PATTERNS AND PROCESSES OF DIVERGENCE IN AFRICAN MARINE SYSTEMS:

IMPLICATIONS FOR CONSERVATION AND MANAGEMENT

by

Amy Jane Elizabeth Healey

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

2019

Aberystwyth University

Institute of Biological Environmental and Rural Sciences

DECLARATION OF WORK

I hereby confirm that the work presented in this thesis was conducted by myself under the supervision of Professor Paul Shaw, Dr. Niall McKeown and Dr. Jim Provan. However Chapter 5 incorporate some microsatellite and/or mtDNA datasets produced in collaboration with Dr. Gavin Gouws of the South African Institute of Aquatic Biodiversity and Dr. Amy Taylor.

Name: Amy Healey

Date: 29/09/2019

ACKNOWLEDGMENTS

I'm so thankful to all the people who supervised me and gave me the opportunity to do this PhD. Thanks to Professor Paul Shaw for all your advice throughout the PhD, I've learnt so much from being in the Shaw lab and I'm so thankful that I had that opportunity. Also I couldn't be more grateful to Dr. Niall McKeown, not only for his top notch rants and funny stories, but also for helping me so much throughout this PhD and imparting just a little bit of his crazy genetics knowledge to me. I wouldn't have got through this PhD without his supervision, guidance, pep talks, positivity and ability to deal with crying girls! I'd also like to thank my second supervisor Dr. Jim Provan for his guidance. I'd also like to thank Professor Warwick Sauer and Dr. Warren Potts at Rhodes University for being my South African supervisors.

There's also a huge amount of people I need to thank for helping me get samples. Firstly Dr. Gavin Gouws from the South African Institute of Aquatic Biodiversity was an amazing person to collaborate with on the eastern African portion of my PhD, providing countless samples and sharing data with me, all of which I am so appreciative of. Also thank you so much to Dr. Francis Nunoo and his students from the University of Ghana for helping me out with my Ghanaian fieldwork, Catalina Perales-Raya and her team for getting me cuttlefish samples from Guinea-Bissau, David Tully and David Stokes from the Marine Institute in Galway for collecting cuttlefish from the Celtic Sea for me, Chanelle de Beer for collecting my Angolan cuttlefish samples and also Gancalo Silva for being so lovely and collecting my Portuguese cuttlefish samples. I'd also like to thank Rachel Mullins for depriving her penguins a few horse mackerel and sending me over some samples and also for being a wonderful host and an amazing friend when I was in Grahamstown. Finally a gigantic thank you to Mathew Farthing for helping out with my South African fieldwork (and also getting me a load of samples from Angola), this guy is the most patient person at trying to teach a girl how to fish and has such a crazy amount of fish knowledge, this trip was honestly one of the highlights of my PhD and Matt was such an integral part of making it so enjoyable. Also a huge thank you to all of Matt's friends and family in South Africa for always being such amazing hosts and throwing the best braais!

Thank you to all the friends I've made in Aberystwyth. Thanks to Max Blake and Mike Gwilliam for answering all of my silly questions and being so helpful when I first started in the lab. In particular I'm so happy I met all the lovely and amazingly clever ladies in my office (Ilze, Rowan & Sophie) and all the lovely people around IBERS, they've all helped me to keep a little bit of my sanity throughout this PhD, have put up with me being the grumpiest person and all appreciate a bottle of wine as much as I do. I'm also so thankful to all my home friends, for sticking with me even though I've had increasingly less free time to come home and be a good friend, the increasingly few and far between trips home have been the most lovely thing and have really helped me keep my sanity and get a little perspective! Thank you also to the Hannah's for always being at the end of the phone.

Thank you so much to my Mum and Dad for getting me through university and for always being there for me, even when I was super grumpy every time you asked how my thesis was coming along! Also

thank you to my grandparents, although none of them managed to make it to the end of this PhD and will never have had the opportunity to pretend to understand and be excited over this thesis, knowing how proud you were of me really helped me to get though the tough times.

Thank you to my beautiful dog Darwin for taking me out for lots of de-stress walkies, being a complete weirdo which always made me smile and always being there for a reluctant cuddle when I needed one. Finally thank you so much to my wonderful other half Dave for putting up with me and being a wonderful housewife/ field assistant while I've been doing this PhD. Thank you so much for always being up for getting dragged to a random African country and for following me around countless smelly fish markets and holding my Eppendorfs! Having you as a constant throughout my time in Aber has meant the world to me and this PhD would have been so much harder without all your support.

ABSTRACT

Through investigating the spatial distribution of genetic diversity and phylogeographic partitioning, this thesis aimed to assess the roles of ecological, environmental and climatic factors in shaping microevolution, adaptation and speciation in a number of marine taxa distributed across the African coastline, with a particular focus on the eastern Atlantic and the South West Indian Ocean (SWIO). Data for a number of taxa (Chapter 2.2, 2.3, 3.2) adds to evidence that the Benguela Upwelling System (BUS) represents a long standing biogeographic boundary that has fundamentally shaped divergence between Atlantic and Indian Ocean gene pools. However, evidence of historical permeability permitting historical gene flow events from South Africa into southern Angola, was reported for several species (Chapter 2.3, 3.2). Most strikingly among Sepia species (Chapter 2.3) this permeability has facilitated complete species hybridisation in Angolan waters. The data also revealed the novel genetic composition for many fish groups along the west and north African coasts and in doing so highlights this area as a biodiversity blind spot and suture zone for divergent genetic lineages (Chapter 4). In contrast to the eastern Atlantic the SWIO was shown to be a high gene flow region, with a number of fish exhibiting high connectivity throughout the region (Chapter 5.1, 5.2), superimposed upon which was a pattern of chaotic genetic patchiness due to stochastic recruitment heterogeneity. Phylogenetic analysis also revealed that two commercially important and hitherto considered distinctive Lethrinus species in the SWIO each comprised two cryptic species (Chapter 5.1, 5.2). From an evo-to-eco perspective the research provides evidence of non-neutral evolution of mitochondrial DNA in Sepia species (Chapter 2.3) and unusually slow mutation rates in horse mackerel (Chapter 4). By resolving patterns and processes from an ecological to evolutionary perspective this research also provides resources and information pertinent to the management and future study of marine biodiversity in both understudied regions.

TABLE OF CONTENTS

CHAPTER 1

Introduction

CHAPTER 1

Introduction

Speciation is a key theme in evolutionary biology, with species representing the fundamental unit in biology. Despite the importance of the species concept, there remains a great deal of contention over the processes that drive speciation, as well as where on the continuum of diversification two evolutionary units can be said to represent discrete species (Panhuis *et al*. 2001; Schluter 2001; Turelli *et al*. 2001; Coyne & Orr 2004; Ridley 2004). In the marine environment the challenges of disentangling speciation mechanisms are further heightened, with high dispersal potential (Roberts 1997; Craig *et al.* 2007) combined with the scarcity of discernible barriers to gene flow (Palumbi 1994) resulting in the assumption that marine populations are demographically open and genetically homogeneous throughout their geographical distributions (Cowen *et al.* 2000; Féral 2002; Craig *et al.* 2007; Horne *et al.* 2008), slowing the processes that drive cladogenesis in marine taxa. Speciation in the marine environment has therefore been considered to be facilitated by a gradual accumulation of genetic incompatibility amongst allopatric populations (Mayr 1954). Despite this assumption of widespread homogeneity, oceans harbour substantial biodiversity, particularly in regions such as the Indo-Pacific which harbour high frequencies of closely related and sympatrically distributed species (Kay & Palumbi 1987; Palumbi 1992). Similarly, in many marine species which conform to assumptions of being highly dispersive, considerable evolutionary diversification has been detected even in the absence of obvious spatial isolating mechanisms (e.g. Palumbi 1992, 1994; Geyer and Palumbi 2003; Zigler and Lessios 2004).

Although numerous theoretical models and classification systems exist to describe speciation processes (see: Templeton 1981, 1996; Kirkpatrick & Ravigné 2002; Richards *et al.* 2018), speciation mechanisms are most commonly characterised based upon the geographical distribution of the diverging populations (Butlin *et al.* 2008), and in particular where they sit on the scale from allopatry to sympatry. In the marine environment allopatric speciation is recognised as the most common mode of speciation (Barraclough & Vogler 2000). However the proliferation of speciation events attributed to this model could reflect the early prominence given to this mode of speciation (e.g. Mayr 1942), in addition to the relative ease with which evolutionary scenarios can be aligned with models of allopatric speciation. The fundamental principles behind allopatric speciation are that when populations become geographically separated, either by chance (genetic drift) or through natural selection, they will diverge both genetically and phenotypically, culminating in reproductive isolation. However complete allopatry is thought to be realised rarely, particularly for marine taxa that have been shown to have the capability for migration across major biogeographic barriers (e.g. the eastern Pacific

barrier: Rosenblatt and Waples 1986; Lessios and Robertson 2006). As such, parapatric speciation is said to occur where taxa have speciated despite a small proportion of range overlap among populations, though empirical evidence of this in wild populations is scarce (Niemiller *et al*. 2008; Nosil *et al*. 2002; Coyne & Orr 2004).

Another derivative of allopatric speciation is peripatric speciation, sometimes referred to as founder events, where reproductive isolation occurs through the separation of a few individuals from a larger ancestral population (Templeton 2008). Although disentangling speciation via founder events from other modes of allopatric speciation has proven difficult, with limited definitive empirical support (Coyne & Orr 2004), the importance of this mode of speciation in evolutionary theory has long been accepted (e.g. Mayr 1954; Carson 1968). The most convincing cases of founder event divergence in marine species are recorded in littorinid snails (Knight *et al.* 1987) and laboratory cultures of polychaete worms (Weinburg *et al.* 1992).

On the opposite end of the scale from allopatry is sympatric speciation, where reproductive isolation of two groups occurs within a single spatial / temporal, and previously panmictic population, in the absence of any geographic isolation or introgression (Coyne & Orr 2004; Richards *et al.* 2018). For terrestrial organisms complete sympatry is rare (Endler 1977), however for highly dispersive marine taxa inhabiting the comparatively more homogenous marine environment, speciation in sympatry is a far more conceivable concept. However there have been very few (if any) convincing cases made for complete sympatric speciation in the marine environment, due in part to the difficulty of ruling out historical allopatric scenarios and/or introgression (reviewed in Richards *et al.* 2018).

Although these traditional models of speciation in allopatry and sympatry remain the most ubiquitous definitions employed in evolutionary biology, they face numerous criticisms for their simplicity and failure to integrate the dynamic and interacting processes driving speciation (Grant & Grant 1997; Schulter 2001; Rundle & Schulter 2004; Rundle & Nosil 2005; Templeton 2008). Accordingly, growing evidence suggests that cladogenesis for many organisms does not occur solely by one of these simplistic geographically based models, but is often initiated in allopatry followed by a subsequent sympatric stage (Feder *et al.* 2003; Albert & Schluter 2004; Rundle & Schluter 2004). The recognition that multiple mechanisms can interact along the speciation continuum has fostered a shift to a process-based classification of speciation mechanisms. Templeton (1981) suggested broadening the categories in which speciation events are defined, so that they can loosely be categorised into “divergence” (where

an ancestral population is split and forms reproductively isolated populations) and "transilience" (where some event initiates the speciation process, e.g. hybridization) mechanisms, and acknowledged that both mechanisms may interact throughout the speciation process (Templeton 1981, 1996). Kirkpatrick & Ravigné (2002) suggested that rather than defining speciation mechanistically, a requirement-based approach should be employed that integrates the commonalities amongst traditional speciation models. Five requirements were proposed: 1) the presence of a source of disruptive selection; 2) some mechanism of pre-zygotic isolation; 3) some interaction between the selective force and the isolating mechanisms; 4) a genetic basis to the increase in isolation; and 5) that there is some driver that initiates divergence. Building upon this, and highlighting the multi-dimensionality of the speciation process is the principle of 'speciation routes' (Dieckmann *et al.* 2004). Although this principle lacks the incorporation of the mechanisms that may underlie divergence, it highlights that the speciation event should follow a 3-step process: 1) a random mating population; 2) external environmental changes, genetic drift and selection; and 3) complete reproductive isolation. However, unlike traditional models, at any point on these speciation routes the speciation process can be influenced by spatial, mating and/or ecological differentiation.

Models such as those of Dieckmann *et al.* (2004) acknowledge the complexities of the speciation process and highlight the frequency with which interactions of various mechanisms drive cladogenesis (Templeton 2008). As such it has been proposed that rather than regarding speciation as a linear process, the complex interactions of the various mechanisms should be viewed multi-dimensionally. Although divergence may be initiated through allopatric, sympatric or parapatric processes these do not occur in the absence of a diverse array of other evolutionary influences inherent to the biological, ecological or geographic context of that specific event in the speciation process (Rundle & Nosil 2005; Butlin *et al.* 2008; Templeton 2008). Nonetheless across multiple species concordant physical environmental features are often recognised as initiating the speciation process (e.g. in the marine environment: the rise of the Panamanian Isthmus – Bermingham *et al.* 1997; Knowlton & Weigt 1998; Marko 2002, and the closure of the Tethys Sea – Hrbek & Meyer 2002; Nohara *et al.* 2004; Teske *et al*. 2004; George 2006).

1.1 Contemporary circulation and oceanography of the eastern Atlantic Ocean, and its connection with other ocean basins (the Mediterranean Sea and Indian Ocean)

The contemporary physical, chemical and oceanographic features of a marine region can have a profound impact on the spatial distribution, dispersal pathways and ultimately the diversification of the biota inhabiting it (Perez-Losada *et al.* 2002; Lessios *et al.* 2003; Teske *et al.*

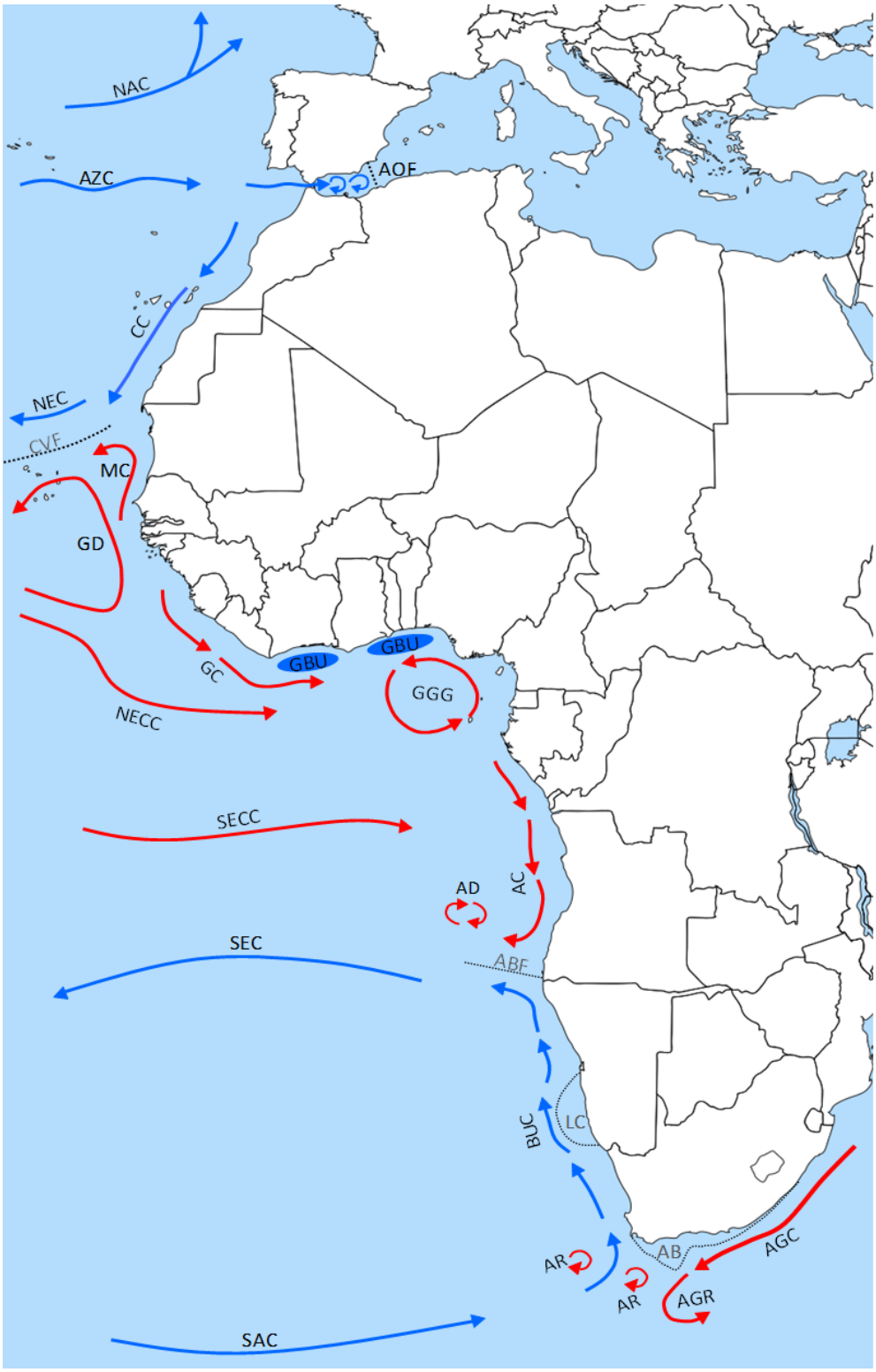


Figure 1.1: Major circulatory pathways of the eastern Atlantic. Currents connecting the east and west Atlantic: SAC=South Atlantic current, SEC= South Equatorial Current, SECC= South Equatorial counter current, NEC= North Equatorial Current, NECC= North Equatorial counter current, NAC=North Atlantic Current. Coastal Currents: AGC= Agulhas current, AGR= Agulhas Retrofelction, AR=Agulhas ring, BUC=Benguela Upwelling current, AC= Angola Current,. AD= Angola Dome, GGG= Gulf of Guinea Gyre, GBU=Guinea Basin seasonal upwelling, GD= Guinea Dome, GC= Guinea Current, MC= Mauritanian Current, CC= Canary Current, AZC= Azores Current. Frontal zones: ABF= Angola Benguela Frontal zone (rough positioning as migrates seasonally), CVF= Cape Verde Front. AOF=Almeria Oran Front. Other Oceanographic features: AB= Agulhas Bank, LC=Luderitz Upwelling cell

2008; Henriques *et al.* 2012). The eastern Atlantic Ocean harbours a complex circulatory system (Fig. 1.1) with a broad band of equatorial tropical water bound to the north (Canary Current Upwelling System-CCUS) and south (Benguela Upwelling System-BUS) by cool upwelling currents, and as such the region is characterised by complex oceanographic systems and boundaries exhibiting steep temperature, nutrient and salinity gradients (Longhurst 1962; Hutchings *et al.*. 2009; El-Geiziry & Bryden 2010). Furthermore, the South African coastline represents the sole point of contact between the Atlantic and Indo-Pacific Oceans for warm temperate and tropical taxa. Consequently the position of the cool Benguela Upwelling System (BUS) along the western South African, Namibian and southern Angolan coastlines may serve to limit the connection between the biota of these two ocean basins, or at least shape the evolutionary history of species and species groups distributed across this ocean boundary.

Below, the contemporary oceanographic conditions and circulation systems of the eastern Atlantic and its connection with other oceanic basins are described, and this expanse of ocean is partitioned based upon the major biogeographic provinces and their boundaries as identified by Spalding *et al.* (2007; Fig. 1.2): 1) the Mediterranean-northeast Atlantic transition; 2) the Canary Current Upwelling System (from the Gulf of Cadiz to Mauritania); 3) the tropical coast of west Africa (from northern Senegal to Angola); and 4) the Benguela Upwelling system (from southern Angola to the western South African coast).

1.1.1 The Mediterranean-northeast Atlantic transition

The Mediterranean Sea is a large semi-enclosed basin, an almost entirely isolated system if it was not for the exchange of water, solvents, and heat with the northeast Atlantic through the Strait of Gibraltar at its western end (Robinson *et al.* 2001). Through this narrow passageway (12.9 km wide, 284m deep), waters are exchanged between the Atlantic and Mediterranean basins, with the flow of less dense Atlantic waters into the Mediterranean concentrated at the surface, whereas the counter-current exchange from the Mediterranean into the Atlantic ocean occurs far below the sea surface (Millot 1999). Within the Alboran Sea the Atlantic inflow forms two quasi-permanent anti-cyclonic gyres, which converge with the western Mediterranean waters, generating the Almeria-Oran Front (see Fig. 1.1) , across which a marked temperature (1.4°C) and salinity (2ppt) gradient is observed (Tintore *et al.* 1988).

Within the Mediterranean basin conditions are characterised by a narrow continental shelf, bound to the north by mountainous Europe, the south by Africa, with the eastern Mediterranean enclosed by the Asian continent (Millot & Taupier-Letage 2005). The Sicilian-Tunisian sill, which reaches a depth of 400m, forms a divide between the east and west Mediterrane-

an, essentially partitioning this sea into two large and distinct basins (Robinson *et al.* 2001). The Alboran Balearic and Tyrrhenian seas make up the western Mediterranean basin (Zavatarelli & Mellor 1995; El-Geziry & Bryden 2010), whilst the features of the eastern basin are far more complex, owing to the fact it is composed of four sub-basins (the Ionian, Levantine, Adriatic and Aegean seas), each with its own distinct environmental conditions.

Being one of the most studied water basins globally, across a range of disciplines, there is a wealth of knowledge regarding the circulatory pathways of the Mediterranean (reviewed in El -Geiziry & Bryden 2010). The interaction between the series of basins and sub-basins that make up the Mediterranean Sea generates a complex and variable circulatory system, with circulation of the entire basin driven predominantly by water exchange through the collection of straits connecting these basins and sub-basins (Robinson *et al.* 2001). However as this thesis focusses primarily on the inter-oceanic connection between the Atlantic and Mediterranean basins, further discussion will be focussed on this area (but see: Robinson *et al.* 2001; Millot & Taupier-Letage 2005; El-Geiziry & Bryden 2010).

1.1.2 The Canary Current Upwelling System

The Canary Current Upwelling System (CCUS) extends from the Strait of Gibraltar along the coasts of Morocco and Western Sahara where it meets the Cape Verde Front (corresponding to ecoregion 10 in Fig. 1.2), a frontal zone that separates the cool upwelled waters of the Canary Current and the tropical waters of the Gulf of Guinea Province. The CCUS is dominated by the Canary Current, an eastern boundary upwelling system which alongside the BUS is one of the 4 major upwelling systems globally. As a result of the CCUS, sea surface temperatures are relatively cool along the north-west African coast (<21°C) (Wooster *et al.* 1976; Mix *et al.* 1986).

The CCUS is dominated by three main currents: the Azores current, the Canary Current and the North Equatorial Current (NEC). The Azores current originates as a branch of the Gulf stream, flowing south-east until it reaches the continental north west African coast, where it bifurcates, with small fractions passing into the Mediterranean basin (Candela 2001), or northwards feeding the Iberian current (Haynes *et al.* 1993), whilst the main flow recirculates south, feeding either the Canary current or the westward flowing NEC (Maillard & Käse 1989; Klein & Siedler 1989; Alves & Colin de Verdière 1999; Knoll *et al.* 2002). The Canary Current begins at the Gulf of Cadiz, and flows south as far as Cape Blanc. The presence of the Canary Islands within the main flow of the Canary Current elicits substantial and complex mesoscale activity (Barton *et al.* 1998), making the CCUS unique from other major upwelling currents. Unlike the Lüderitz cell of the Benguela Upwelling System the CCUS lacks a predominant and

perennial upwelling cell, and is instead comprised of many seasonally influenced regions of intensified upwelling. Of these, the upwelling cells found off Cape Ghir (31-32°N) and Cape Blanc (19-21°N) are arguably the most distinct (van Camp *et al.* 1991), with high intensity and persistent upwelling observed, although upwelling strengthens during boreal summer at Cape Ghir (Hagen *et al.* 1996; Hernandez-Guerra & Nykjaer 1997; Barton *et al.* 1998) and winter at Cape Blanc (Longhurst 2006). Of the other upwelling cells a high degree of seasonal variability is observed, due to migration of the trade wind system; though in general in the northern segment (above 25°N) of the CCUS upwelling is strongest in boreal summer with more continuous upwelling observed in the central region (20-25°N) and intensified upwelling in the southern portion (below 20°N) of the CCUS observed during the boreal winter (Mittelstaedt 1991; Van Camp *et al.* 1991; Hagen *et al.* 1996; Barton *et al.* 1998; Navarro-Perez & Barton 2000). At around 15-20°N the southward flowing Canary Current converges with the northward flowing Mauritanian and Poleward undercurrent, forming the Cape Verde Front (Zenk *et al.* 1991; Arhan *et al.* 1994) which stretches from the west African coast to the Cape Verde islands.

1.1.3 Tropical west Africa

The tropical west African coast (Gulf of Guinea Province in Spalding *et al.* 2007; Fig. 1.2), extends south from the Cape Verde Front in northern Mauritania, throughout the Gulf of Guinea and as far south as southern Angola. Compared to the other regions included in this investigation the tropical west coast of Africa has been relatively poorly characterised, not only in regards to its oceanography and geological history, but also the ecology and evolutionary history of the species that inhabit it. Despite this, available investigations highlight a high degree of endemism within the tropical eastern Atlantic in reef fish (30%; Floeter *et al.* 2008), tunicates (31% Naranjo *et al.* 1998) and gastropods (36%; Garcia & Bertsch 2009), such that a more comprehensive characterisation of how the oceanography and geological history of the tropical west African coast have shaped these patterns is required. The tropical eastern Atlantic is affected by 7 predominant current systems: the BUS, The South Equatorial current (SEC), the North and South Equatorial Counter currents (NECC/SECC), The CCUS and the Guinean and Angolan currents (Longhurst 1962), and their interaction with the seasonal migration of the Inter Tropical Convergence Zone (ITCZ) in the atmosphere.

The ITCZ is a zonal band that loosely follows the trajectory of the earth's meteorological equator, and marks the location in which northern and southern trade winds converge (Lietzke *et al.* 2001; Henke *et al.* 2012; Schneider *et al.* 2014). The ICTZ migrates seasonally between 5°S and 15°N (Waliser & Jiang 2014; Waliser & Gautier 1993), however north-south asymmetry in this migration has been observed, with occurrence of the ITCZ near to the equator and the

southern hemisphere only in the boreal spring (Chiang *et al.* 2002). This latitudinal shift in the ITCZ has a profound impact on the ecological and biochemical properties of the tropical west African coast, influencing numerous physical processes, including coastal upwelling, riverine inputs and vertical mixing of the thermocline (Folland *et al.* 2002), which have knock-on effects on sea surface temperature (SST), productivity and ultimately the biodiversity of the tropical eastern Atlantic.

The tropical west African coast can be subdivided into 3 regions. The first lies in the far north, across the Mauritanian and Senegalese coasts (Sahelian Upwelling ecoregion in Fig. 1.2) and is the southerly extension of the CCUS. However as this region is not influenced by the main upwelling cells of the CCUS this southern region is dominated by two warm currents, the north flowing Mauritanian current, an extension of the North Atlantic Central Current, and the anticyclonic Guinea Dome (Siedler *et al.* 1992; Pelegrí & Peña-Izquierdo 2015). The other two regions that make up the tropical west African coast can be loosely subdivided by the ITCZ: north of this convergence zone, the Guinea Coast region extends from northern Guinea-Bissau to the Bight of Biafra in Nigeria; whilst south of the ITCZ the central west African coast extends from Nigeria south to southern Angola.

The Guinea coast region can be further subdivided into the western and central portions of the Gulf of Guinea. The western portion is dominated by the Guinea Current, which carries warm (~27° C) water along the west African coast between 2°N and 5°N, fed predominantly from the NECC (Ingham 1970; Henin *et al.* 1986; Djakoure *et al.* 2017). The intensity of the Guinea Current peaks during the boreal summer (May-September), with a minimum observed during winter months (November-February) (Boisvert 1967; Ingham 1970; Richardson and Philander 1987). The central portion of the Gulf of Guinea is dominated by seasonal upwelling, with cool water penetrating the sea surface from June to September (Houghton & Colin 1986; Hardman-Mountford & McGlade 2003; Okumura & Xie 2004). This upwelling process is variable amongst years, likely influenced by the position of the ITCZ (Peterson & Stramma 1991), but the regions that most frequently exhibit significant upwelling are those downstream of Cape Three Points in Ghana (4° 44”N, 2° 5”W) and Cape Palmas in Liberia (4° 22”N, 7° 43”W) (Lemasson & Rebert 1973; Koranteng & McGlade 2001; Hardman-Mountford & McGlade 2003; Djakoure *et al.* 2017).

Circulation patterns along the central African coast, which extends from Nigeria to southern Angola, are equivalent to those of the Guinea region, at least for the northern portion (north of the Congo river), and is dominated by the substantially weakened southward flowing Guinea Current. However the eastern section of the Gulf of Guinea Gyre forms the Angola current,

a fast, warm (24-30°C) and narrow current (Moroshkin *et al.* 1970; Lass *et al.* 2000) that flows south, becoming gradually cooler until it reaches the Angola-Benguela Front, where it is deflected offshore.

The sub-tropical coasts of Angola are bounded to the north by the Congo River freshwater flume and to the south by the Angola-Benguela Front (ABF) (Hutchings *et al.* 2009), and represent a complex environment. The ABF forms at the convergence of the Benguela Current and the southward flowing Angola Current, forming the northern boundary of the Benguela upwelling region at around 16^{o}S (Shillington *et al.* 2006; Hutchings *et al.* 2009). The Angolan subtropical zone is characterised by pronounced seasonal variability in oceanic conditions, in particular sea surface temperature. The precise locality of the ABF is associated with the position of the South Atlantic Anticyclone and as such its position is altered seasonally: displacing southward to around 23^{o}S during austral summer, resulting in an influx of warm equatorial water; and shifting north to around 12^{o}S in austral winter allowing for the influx of cool south Atlantic water into this region (Shannon 1985; Meeuwis & Lutjeharms 1990).

1.1.4 The Benguela Upwelling System & Atlantic-Indian Ocean boundary

The Benguela region (defined here as the coastal area from southern Angola and throughout Namibia to western South Africa), like the CCUS, is dominated by one of the five great continental margin upwelling systems, and is characterised by Ekman driven upwelling of cool nutrient rich waters. However unlike other upwelling systems globally, the Benguela Upwelling System (BUS) is unique in being bound to both the north and south by the tropical waters of the Angola Current and Agulhas Current respectively (Boyer *et al.* 2000), and as such provides an interesting model for connectivity and diversification of tropical species across it. The BUS can be divided into two biologically, physically and chemically differentiated subsystems (the northern and southern Benguela), partitioned by the perennial Lüderitz upwelling cell in central Namibia (Shannon 1985; Hutchings *et al.* 2009).

The northern Benguela subsystem extends from southern Angola to central Namibia. Much like the Angolan subtropical zone, the oceanic conditions of this region are dictated by the location of the Angola-Benguela Front (ABF), namely the dynamic balance between the flow of cold upwelled waters from the south against warm tropical waters from the north (Boyer *et al.* 2000; Hutchings *et al.* 2009). Consequently the southern Angolan and northern Namibian coasts are characterised by sharp gradients in SST (Shannon 1985), as well as an area of water with low oxygen levels across the northern Namibian shelf (Hutchings *et al.* 2009). The central Namibian coast is defined by largely homogenous environmental conditions, dominated by shallow and low intensity upwelling due to the combination of favourable winds and a

shallow continental shelf (Boyd 1987; Boyer *et al.* 2000). Where the continental shelf steepens at Palgrave point and Conception Bay an increase in upwelling intensity has been observed (Boyer *et al.* 2000). The southern boundary of the northern Benguela subsystem is thought to fall between 21°S and 23°S where the presence of a semi-enclosed circulation system (Barange & Boyd 1992) forms a semi-permanent convergence zone between the waters of the Lüderitz upwelling cell and the northern Benguela subsystem.

The Lüderitz upwelling cell is the strongest upwelling cell of the BUS (Shannon 1985; Lutjeharms & Meeuwis 1987; Summerhayes *et al.* 1995; Hutchings *et al.* 2009), and is one of the most persistent and intense regions of upwelling globally (Boyd *et al.* 1987). This upwelling cell sees water fed predominantly by the South Atlantic Current rise from depths of 250-350m (Shannon *et al.* 1989; Boyer *et al.* 2000). The unprecedented strength of upwelling in this area can be attributed to several factors, namely strong and persistent southerly winds and the presence of deep and narrow continental margins (Boyd *et al.* 1987; Boyer *et al.* 2000). Furthermore, upwelling strength is correlated with the South Atlantic cyclone, exhibiting cyclic annual and decadal variability (Sakko 1998). Consequently the Lüderitz upwelling area is characterised by high levels of offshore advection, resulting in a predominantly north westerly flowing band of highly turbulent upwelled water, with negligible vertical stratification and markedly low phytoplankton levels (Agenbag & Shannon 1988; Hutchings *et al.* 2009).

South of the Lüderitz upwelling cell the southern Benguela subsystem extends across southern Namibia and along the western coast of South Africa as far as the Agulhas Bank. Across the southern Benguela subsystem upwelling intensity is spatially and temporally variable, peaking in late austral summer to early autumn. Cold highly productive water converges upon the Agulhas Bank where it is met by the warm saline Indian Ocean waters of the Agulhas Current (Hutchings *et al.* 2009). Similarly to the northern Benguela subsystem, the southern Benguela subsystem is characterised by low oxygen levels thought to be a product of the decay of phytoplankton blooms (Pitcher & Weeks 2006), though oxygen depletion is far less pronounced in the southern than in the northern Benguela subsystem (Hutchings *et al.* 2009).

1.2 Climatic and geological history of the eastern Atlantic Ocean

In order to predict how a species may respond to future climatic changes an understanding of their interaction with past climatic events is essential. A species response to past climatic events, in particular the Pleistocene glacial/interglacial cycles (Lecomte *et al.* 2004; Grant & Bowen 2006; Janko *et al.* 2007), provide an insight not only into the causal factors underly-

Table 1.1: Summary of major climatic and geological events and their influence on the Mediterranean-north east Atlantic, Canary Current Upwelling System, tropical west Africa and Benguela Upwelling regions of the eastern Atlantic Ocean, from the early Miocene epoch (~23 MYA) to the end of the Last Glacial Maximum (~0.1 MYA).

		Global	Med-NE Atlantic	Canary Current	Tropical W Africa	Benguela region
Early Miocene	23-14.9MYA	Global climate significantly warmer than today.	Relatively stable environment connected to the Indian Ocean by the Tethys Sea, and as such dominated by Indian Ocean taxa.	Increase in upwelling along the north west African coast, representing the early stages of the Canary Current upwelling system.	Early Miocene - less pronounced gradient in SST from the tropics to the poles.	Southeast Atlantic upwelling not established yet.
Mid Miocene	~14.8-8MYA	Mid Miocene climatic transition. Uplift of the Panamanian Isthmus and expansion of Antarctic Ice sheets resulted in cooling of sea surface and deep water.	Closure of the Tethys Sea (16-14MYA) isolated the Mediterranean from the Indian Ocean.	Upwelling intensity strengthened	Expansion of glacial ice sheets at poles steepened the thermal gradient between the tropics and the poles	The mid Miocene climatic transition, strengthened oceanic circulation systems, giving rise to the Benguela Upwelling system.
Late-Miocene/early Pliocene	8.7-5MYA	The onset of pronounced Antarctic glacial-interglacial cycles	During the Messinian salinity crisis (5.9-5.3MYA) the Mediterranean basin was separated from the Atlantic Ocean and became near desiccated. Tectonic uplifting (~5MYA) changed sea levels and flooded the Med basin with Atlantic waters.	Upwelling intensity began to drop in the late Miocene.	Formation of the seasonal Gulf of Guinea upwelling cells.	Glacial/ interglacial cycles drove shifts in the trajectory of the Benguela Current, northwards during Antarctic glaciation. At ~6MYA the Benguela Current System became fully established.
Mid Pliocene	4.9-3MYA	Pliocene warm period, characterised by ~3degC higher global surface temperatures and 10-20m higher sea levels than today	Climate conditions improved, increasing marine production in the Mediterranean (~3MYA)	Upwelling intensity was sluggish	~4.92MYA Establishment of the Guinea Current, ~4.33MYA establishment of the modern seasonal migration cycle of the ITCZ	Upwelling intensity was sluggish.
Late Pliocene-early Pleistocene	3-2MYA	Significant intensification of northern hemisphere glaciation (~2.75MYA).	Atlantic waters began to flow regularly into the Mediterranean (2.5-1.8MYA), forming the flow regime observed presently.	Intensification of upwelling in the CC, significantly reducing SST	SST in the GG upwelling system began to decrease (~3MYA).	Upwelling intensification (2.9-1.2MYA). Establishment of the Luderitz upwelling cell, giving rise to the present day oceanographic features of the BUS.
Mid-late Pleistocene	~1.9-0.1	The Pleistocene was characterised by 100,000 year glacial -interglacial cycles, resulting in global fluctuations in circulation, SST and sea levels.	Changing sea levels, altered/ restricted flow regimes between the west and east Mediterranean, as well as between the Mediterranean and Atlantic.	Intensification of upwelling accompanied by a southerly shift in the flow of the CC during glacial maxima.	Arid, cooler and more upwelling in the GG during glacial maxima.	Glaciations led to southern displacement of ABF and fluctuations in upwelling intensity. Lowered sea levels exposed the Agulhas Bank, elongating the South African coast by around 200km.

ing contemporary distributional limits of a species, but also on rates of migration and population resilience (Willis and Niklas 2004). However, any inference regarding a marine species response to historical climate change requires an understanding of the drivers and extent of historical climatic events and their impact on the oceanic environment. Accordingly, the post-Paleogene impact of large scale changes in climate and geology across the study area (Mediterranean-north east Atlantic, CCUS, tropical eastern Atlantic and the BUS) is reviewed with the aim of providing a brief overview (see Table 1.1) of how these changes influenced the oceanic environment of the eastern Atlantic.

Globally the Miocene epoch (23-6 MYA) saw great changes in oceanographic circulation and was a profoundly important period in the formation of contemporary oceanic circulation patterns; it was in this period that many of the major current systems that dominate the earths oceans today became established. The Miocene was characterised by changing climatic conditions, from warm global temperatures in the early Miocene to a cooler climate by the middle Miocene (Savin *et al.* 1981; Woodruff & Savin 1989; Flower & Kennett 1993; 1994).

The early Miocene was the warmest period in the last 25 million years (Zachos *et al.* 2001), and as such the SST gradient from the equator to the poles was far less pronounced than observed today (Nikolaev 2006; Bruch *et al.* 2007), limiting the potential for thermal barriers to dispersal for marine taxa, at least across tropical equatorial waters. The Miocene appears to be an important period in the formation of eastern Atlantic upwelling currents: the early Miocene saw an increase in upwelling along the northwest African coast, which has been linked to the formation of the upwelling cells that make up the present day CCUS (Emery & Uchupi 1984). Furthermore the Mid-Miocene climatic transition (MMCT; 14.2-12.8 MYA) saw expansion of Antarctic ice sheets, reinforcement of the Antarctic Circumpolar Current and the onset of the Isthmus of Panama uplift, all of which promoted global lowering in both surface and deep water temperatures (Krammer *et al.* 2006; Heinrich *et al.* 2011). These mid-Miocene changes are thought to have driven an increase in the intensity of upwelling in the southeast Atlantic (Diester-Haass *et al.* 2002), giving rise to and firmly establishing the South Atlantic Upwelling cell (the early stages of the Benguela Upwelling System) around 12 MYA (Siesser 1980; Diester-Haass *et al.* 1990; Krammer *et al.* 2006).

The MMCT was also a period of profound change within the Mediterranean basin: the beginning of this period saw the closure of the Tethys Sea (16-14 MYA), which not only stimulated a major reorganisation of oceanic circulation globally (Flower & Kennett 1994; Shevenell *et al.* 2004; Holbourn *et al.* 2013), but also halted water exchange between the Mediterranean Sea and Indian Ocean basins. Around 12 MYA a steep decrease in global carbonate concentrations

(the carbonate crashes) occurred and consequently global ocean productivity plummeted (Jiang *et al.* 2007).

The transition period of the late Miocene to early Pliocene (~7MYA) was characterised by the onset of pronounced glacial/ interglacial cycles, with the increase in glacial ice sheet volume at both poles leading to major sea level regressions and a global reduction in SST. These glacial cycles appear to have had a profound impact on the southeast Atlantic upwelling cell, such that the Benguela current system became fully established around 6MYA. More specifically sedimentary evidence across the south east Atlantic suggest shifts in the trajectory of the Benguela current System (Diester-Haass *et al.* 2002; Diester-Haass *et al.* 1990; Krammer *et al.* 2006), the flow of which appeared to have shifted northwards during the Miocene-Pliocene transition into a location comparable to its contemporary situation (Krammer *et al.* 2006). Similarly the position of the Intertropical Convergence Zone shifted north during the late Miocene and early Pliocene (Hay & Brock 1992; Pisias *et al.* 1995). Fossil records and isotopic data suggest that as a result of this change in the Intertropical Convergence Zone the tropical coast of western Africa was considerably warmer than present (Yemane *et al.* 1985), which may have resulted in the strengthening of the latitudinal SST gradients characteristic of contemporary tropical equatorial conditions (Valentine 1984; White 1986). The onset of glacial-interglacial cycles during the Miocene-Pliocene transition also correlate with the first appearance (~5.65MYA) of the seasonal upwelling cells in the northern Gulf of Guinea (Wagner 2002).

Environmental changes during the early Pliocene had profound impact on the Mediterranean basin and its fauna. Tectonic processes in combination with glacial eustatic changes during the early Pliocene progressively restricted and ultimately isolated the Mediterranean basin from the Atlantic Ocean between 5.59-5.33MYA (Hsü *et al.* 1973; Ryan 1973; Hodell *et al.* 1994; Clauzon *et al.* 1996; Krijgsman *et al.* 1999). As a result the Mediterranean basin became nearly completely desiccated over a period of half a million years (the Messinian Salinity Crisis), driving the majority of pre-existing Mediterranean species (of Indian Ocean origin) to extinction (Penzo *et al.* 1998; Hrbek & Meyer 2003; Huyse *et al.* 2004). The Messinian salinity crisis concluded around 5MYA when tectonic uplifting promoted sea level changes and initiated catastrophic flooding of the Mediterranean basin with Atlantic waters through the newly opened Straits of Gibraltar (Rögl 1999; Almada *et al*. 2001; Domingues *et al.* 2005). Consequently the contemporary biodiversity of the Mediterranean is of comparatively recent origin, and largely the result of colonisation by Atlantic taxa.

The middle Pliocene (4.6-3.2MYA) was the last period in geological history where global climate was significantly warmer (by ~3-8°C) than it is today (Dowsett *et al.* 1992; 1996), particu-

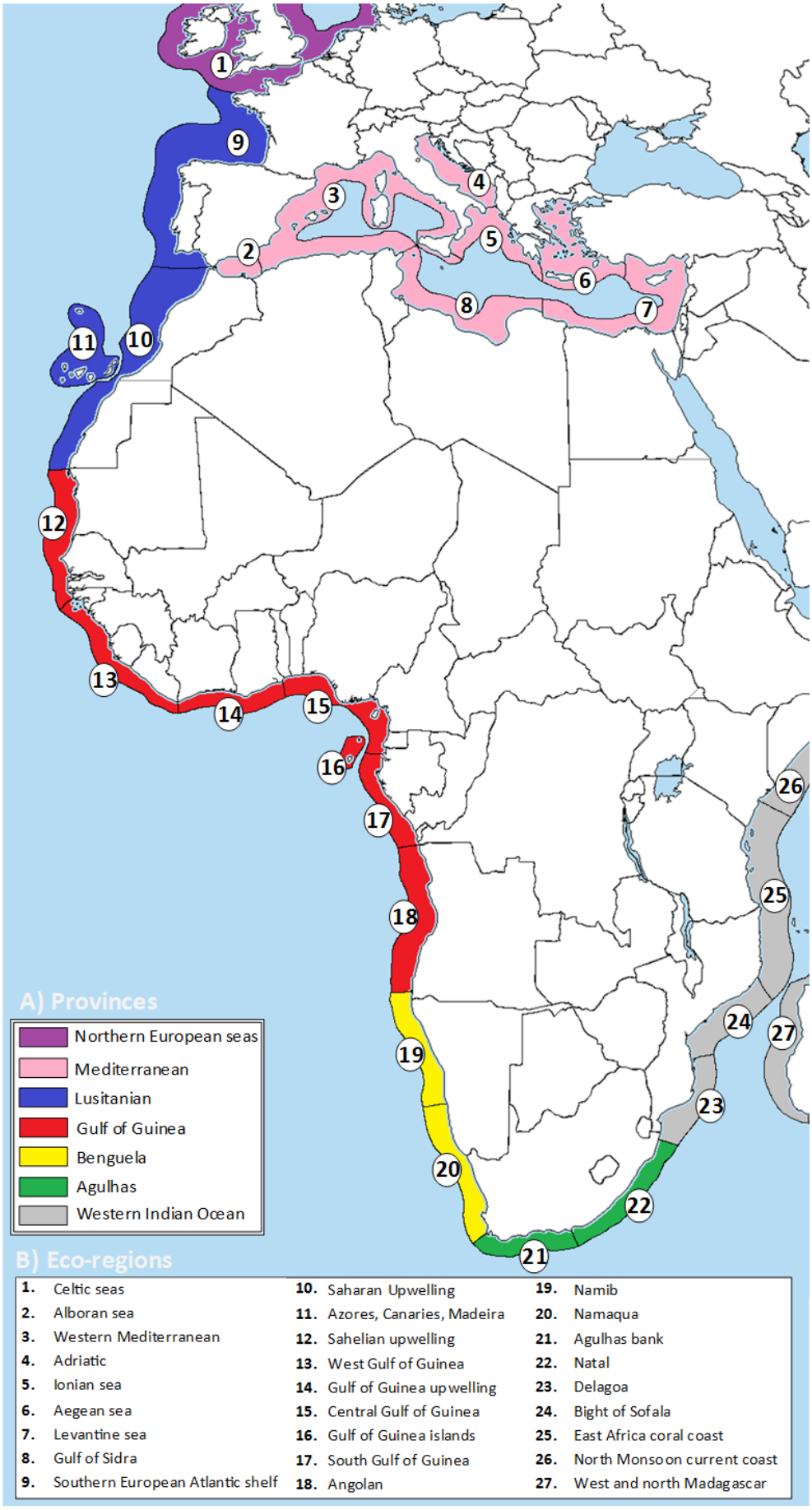


Figure 1.2: Biogeographic A) provinces (colours) and B) ecoregions (numbers) as identified in Spalding *et al.* (2007)

larly in higher latitudes. The warm climate of the middle Pliocene induced a sharp increase in marine production within the Mediterranean basin, such that by 3MYA the environment was habitable again after the Messinian salinity crisis. Within the tropics and unlike higher latitudes, SST was not significantly greater than today (Dowsett *et al.* 1996). However weakening of the thermocline in the equatorial Atlantic around 4.92 MYA in association with the first instance of the NECC penetrating as far east as the Gulf of Guinea led to the establishment of the Guinea Current (Norris 1998). Also around this time (5.3-4.33MYA) sedimentary evidence suggests that the Intertropical Convergence Zone had reached its contemporary position (Hovan 1995, Pisias *et al.* 1995, Billups *et al.* 1999; Wagner 2002). This relocation of the Intertropical Convergence Zone was likely associated with shifting and less intense trade winds, as a result of which Ekman-driven upwelling was lowered in both the Canary and Benguela upwelling systems between 5.3 and 3.2 MYA.

The closure of the Central American Seaway by the Isthmus of Panama (~3MYA), and the consequent production of north Atlantic deep-water during the late Pliocene, initiated a global climatic change from relative warmth in the mid Pliocene to a significantly cooler climate during the mid to late Pleistocene (Hodell & Venz 1992; Marlow *et al.* 2000). This global climate change led to the formation of glacial ice sheets in not only the Antarctic (as was the case in the Pliocene) but also significant northern hemisphere glaciation around 2.6MYA (Maslin *et al.* 1998). This transitionary period between the Pliocene and Pleistocene (3-2MYA) was consequently a time of substantial oceanic cooling, glaciation and eustatic sea level reduction (Siesser 1980). As a result of these eustatic changes, between 2.5 and 1.8MYA (García *et al.* 2009; Llave *et al.* 2007; Rogerson *et al.* 2012; Colleoni *et al.* 2012) the exchange of water from the Atlantic to the Mediterranean increased to a level that is comparable to those observed today, representing the first substantial penetration of Atlantic waters into the Mediterranean since the termination of the Messinian salinity crisis around 5MYA. In both the Benguela (Diester-Haass *et al.* 1992; Dowsett & Willard 1996; Hay & Brock 1992; Marlow *et al.* 2000) and Canary (Marlow *et al.* 2000) upwelling systems late Pliocene/ early Pleistocene cooling saw the intensification of upwelling regimes. In the BUS this intensification of upwelling gave rise to the Lüderitz upwelling cell, and consequently the present oceanographic features of this system (~2.1-1.9MYA; Robinson *et al.* 2002; Krammer *et al.* 2006; Marlow *et al.* 2000).

The mid to late Pleistocene (1.90-0.2MYA) was defined by rapid global cooling and periods of glaciation, punctuated by interglacial cycles (Pisias *et al.* 1981). This time was therefore one of great variability in the marine environment, that advanced far beyond the high latitudes that were directly affected by glacial conditions. Whilst the late Pliocene and early Pleistocene were characterised by low amplitude glacial-interglacial cycles, the early-middle Pleisto-

cene transition (EMPT) (0.9-0.65MYA) saw the intensification and prolongation of these glacial-interglacial cycles, with a shift from cycles lasting around 41,000 years in the late Pliocene and early Pleistocene to the establishment of the Quaternary 100,000 year glacial-interglacial cycle. Furthermore glacial-interglacial cycles shifted from smooth to more abrupt transitions during the EMPT, with more pronounced SST and eustatic fluctuations between glacial and interglacial periods reported post EMPT (Elderfield *et al.* 2012; Rholing *et al.* 2014; Maslin & Brierley 2015).

The pronounced glacial-interglacial cycles of the late Pleistocene drove rapid changes in sea levels (up to 135m below present levels during the LGM; Fairbanks 1989; Lambeck *et al.* 2002), exposing shelf habitats globally. Sedimentary evidence suggests that these eustatic changes had a profound impact on the Mediterranean basin and its connection to the Atlantic Ocean. Falling sea levels during glacial maxima resulted in habitat fragmentation between many of the basins and sub-basins within the Mediterranean (Bargelloni *et al.* 2005), theoretically restricting the potential of species to disperse among these basins. In particular, the depth and width of the Gibraltar passage reduced substantially during glacial maxima, essentially halting the connection between Atlantic and Mediterranean water masses (Patarnello *et al.* 2007). This fragmentation of the Mediterranean basin in combination with sharp changes in the SST and chemical properties of the Mediterranean waters caused localised extinction of Mediterranean taxa as migratory pathways were obstructed, providing limited opportunity to track isotherms (Cunningham, & Collins 1998).

Late Pleistocene glaciations also appear to have had a marked effect on eastern Atlantic upwelling, largely due to the intensification of trade winds during glacial maxima. In the CCUS sedimentary evidence suggests that waters were 8ºC cooler than today and that upwelling intensified, extending the zone of upwelling further offshore and shifting the position of the main upwelling cells 500km further south (Emery & Uchupi 1984). A southern displacement of the BUS was also observed, driven by a concordant southerly shift in the southern Atlantic anticyclone and the subtropical convergence zone (Jahn *et al.* 2003). However, unlike the Canary Current, it is argued that the strength of upwelling in the Benguela region was stronger during inter-glacials (Diester-Haas 1985; Diester-Haas *et al.* 1992; Jahn *et al.* 2003), although numerous other studies cite the converse, with intensification of upwelling reported during glacial maxima (Mix 1986; Dean *et al.* 1984; Oberhänsli 1991). Based on sedimentary evidence Summerhayes *et al.* (1995) suggested that unlike other upwelling systems globally the strength of the Benguela upwelling was independent from the effects of glacial-interglacial cycles. This potential increase in upwelling intensity in combination with the generalised sea

surface cooling in the northern and southern hemispheres resulted in an influx of cool water into the tropical east Atlantic during glacial maxima, with SSTs within the Gulf of Guinea up to 8°C cooler than those observed today (Emery & Uchupi 1984). Furthermore the strength of upwelling cells in the Gulf of Guinea intensified, promoted by the stronger trade winds associated with glacial maxima (Gardner & Hays 1976; Prell *et al.* 1976; Mix 1986) serving to further reduce the SST of this region. During interglacials, however, SST in the tropical eastern Atlantic rose to temperatures far above those observed today (Mix 1986), with this warm water propagating into the south and north east Atlantic, widening the thermal equatorial zone (Mix 1986). The late Pleistocene glacial-interglacial cycles, and consequently the most recent period of global climatic fluctuation, terminated with the last glacial maximum (LGM) around 20,000 years ago, resulting in the eastern Atlantic contemporary oceanographic conditions (section 1.2).

These extended periods of climatic fluctuations, and geological and oceanographic changes, over the last 15 million years not only affected the thermal, circulatory and biochemical properties of the global oceans but also the distribution, abundance and ecology of the species that inhabit them. Molecular genetic (i.e. phylogenetic and phylogeographic) approaches provide the opportunity to associate these past climatic events with the temporal scale of genetic divergence and the partitioning of genetic diversity within and among eastern Atlantic marine species and thus determine the eco-evolutionary response of these taxa to historical environmental perturbations.

1.3 The influence of historical climatic changes and contemporary oceanographic features on the connectivity, demographic history and divergence of eastern Atlantic taxa.

Recurrent fluctuations in the oceanographic features and climatic conditions of the eastern Atlantic (section 1.2) likely had a profound impact on the evolution of the species that inhabit it, through shaping their biology, ecological interactions and distributions (Briggs 1987a,b). Molecular phylogenetics has allowed modes of speciation to be tested directly and relatively precisely, through assessing the signatures of that speciation event that are retained within an organism's genome. When phylogenies are appropriately associated with the biology, ecology and distributions (both historical and contemporary) of a species or species group, powerful inferences can be drawn regarding the drivers of diversification and reproductive isolation (Barraclough *et al.* 1998; Coyne & Orr 1997; Barraclough & Vogler 2000). Furthermore, when dates can be applied to nodes of a phylogeny, speciation events can be associated, albeit indirectly (Rambaut & Bromham 1998), with historical climatic and/or geological events and provide further inferences regarding environmental drivers of diversification.

Through the incorporation of more extensive sampling regimes, within species (or closely related species) and across geographic space, phylogeographic investigations allow the spatial distribution of genetic diversity to be assessed over evolutionary time. As such phylogeography provides the opportunity to more comprehensively test the demographic history of a species / species group, providing evidence regarding the historical processes such as population bottlenecks or expansions that have shaped patterns of speciation (Avise 2000; Nichols 2001).

In the most fundamental sense, biogeographic partitioning defined by patterns of endemism has been shown to be generally concordant with the population genetic and phylogeographic partitioning of species and species groups (Briggs & Bowen 2012). The world oceans have been partitioned into biogeographic regions based on various defining characteristics by a suite of investigations, with the most prominent being those of Ekman (1953), Briggs (1974; 1995), Longhurst (1998), Sherman *et al.* (2006) and Spalding *et al.* (2007). Across the eastern Atlantic coast of Africa these various proposed biogeographic partitions are generally in concordance, at least in regards to the major coastal provinces (however of note is the further subdivision of the Lusitanian province of Spalding and colleagues into a Canary Current and north east Atlantic province for pelagic species by Longhurst). For the present study the biogeographic designations of Spalding *et al.* (2007) are used, in part due to the incorporation of previous biogeographic partitions into these designations (e.g. Briggs 1974; 1995; Sherman *et al.* 2006), but also the finer spatial subdivision provided by the nesting of eco-regions within the major biogeographic province. Based on the divisions proposed by Spalding and colleagues the Mediterranean, northeast Atlantic, west African and southern African coastlines can be divided into 5 provinces: Mediterranean, Lusitanian, Gulf of Guinea, Benguela and Agulhas (see Fig. 1.2). Furthermore, and unique to previous biogeographic investigations, 21 ecoregions defined as "areas of relatively homogenous species composition" (Spalding *et al.* 2007) are nested within the wider provinces, and are indicative of subtle processes and features that have broadly shaped the distributional limits of communities of species and populations.

Not surprisingly, the biogeographic provinces identified by Spalding et al. (2007) can be associated with some of the major oceanographic features as described here in section 1.1, features which may also be associated with diversification and speciation in taxa distributed across the intersection (and within) these major biogeographic provinces. The areas of interest for the present study are: 1) the Mediterranean-northeast Atlantic transition (the intersection between the Mediterranean and Lusitanian provinces); 2) the Canary Current

Upwelling system (the intersection between the Lusitanian and Gulf of Guinea provinces); 3) the tropical west African coast (ecoregions 12 to 18 in Fig.1.2); and 4) the Benguela Upwelling system (the interaction of the Benguela province with the Gulf of Guinea and Agulhas provinces to the north and south respectively). Phylogenetic, phylogeographic and population genetic investigations are contributing to our understanding of how oceanographic features and environmental conditions (and historical changes in these features / conditions) may drive population and species diversification within and among these biogeographic regions.

1.3.1 The Mediterranean-northeast Atlantic transition

The Mediterranean-northeast Atlantic region is arguably one of the most well investigated regions globally, in regards to the diversification of the species that inhabit it (see reviews by Borsa *et al.* 1997 and Patarnello *et al.* 2007). Although the Mediterranean basin harbours substantial endemic biodiversity (Sherman *et al.* 2006; Spalding *et al.* 2007), numerous taxa found within the Mediterranean have distributions that extend, at least to some degree, into the Atlantic (Borsa *et al.* 1997, Bargelloni *et al.* 2003). For many of the species spanning these two oceanic basins the transitional area, whether this be the Gibraltar Strait or the Almeria-Oran Front, has been proposed as a major phylogeographic/population genetic boundary (e.g. Kotoulas *et al.* 1995; Bargelloni *et al.* 2003; Perez-Losada *et al.* 1999; 2002; Zane *et al.* 2000; Catarino *et al.* 2015). Conversely for many other species a distinct lack of genetic partitioning between the Atlantic and Mediterranean has been reported (e.g. Abaunza *et al.* 2004; Stamatis *et al.* 2004). Interestingly these discordant patterns of phylogeographic partitioning between the Mediterranean and Atlantic basins have been reported even between closely related species with similar ecologies and life history characteristics, for which a shared response to a barrier to genetic exchange would be expected. For example, in species groups such as the Sparidae (Bargelloni *et al.* 2003; 2005), lobsters (*Nephrops norvegicus* and *Homarus gammarus* - Stamatis *et al.* 2004; Triantafyllidis *et al.* 2005), anglerfish (*Lophius budegassa* and *L. Pescatorius*- Charrier *et al*. 2006), mackerel (*Scomber japonicus* and *S. scombrus*- Zardoya *et al.* 2004) and large pelagic fish such as *Thunnus thynnus* and *Xiphias gladius* (Alvardo-Bremer *et al.* 2005), stark discordance in evolutionary response to the Atlantic-Mediterranean transition by different but closely related and ecologically similar species has been reported. Often with complete reciprocally monophyletic association of clades between the two basins reported in one species, whilst the other species exhibits signatures more readily aligned with panmixia. This resolved discordance amongst closely related taxa has been attributed to subtle differences in the distribution, population demographics, life history and ecologies of the species, which have driven differences in their responses to historical vicariance.

In a recent review of phylogeographic partitioning across the Mediterranean-Atlantic transition Patarnello *et al.* (2007) identified three generalised modes of genetic structuring of species across this region. The first is panmixia between Atlantic and Mediterranean populations (e.g. *Balaenoptera physalus*, *Diplodus sargus*, *Pagellus bogaraveo*, *Pagrus pagrus* and *Scomber japonicus*), which based on demographic tests could be explained either by high dispersal capacity between the two basins or historical localised extinction of Mediterranean clades with subsequent recolonization from the Atlantic (or *vice versa*). The second, and most common, mode of phylogeographic partitioning was a distinct change in genetic composition of populations between the Atlantic and Mediterranean basins (e.g. *Dentex dentex*, *Dicentrarchus labrax*, *Diplodus puntazzo*, *Hydrobia acuta*, *Lithognathus mormyrus*, *Meganyctiphanes norvegica*, *Pomatoschistus microps*, *P. minutus*, *Spondyliosoma cantharus*, *Sepia officinalis* and *Tursiops truncates*), which was attributed either to contemporary factors such as isolation by distance or ancient separation with divergence still being maintained, most commonly dated to Pleistocene events (Hewitt 1996; Peijnenburg *et al.* 2004; Bargelloni *et al.* 2003; Alvardo-Bremer *et al.* 2005; Chevolot *et al.* 2006; Griffiths *et al.* 2011; Catarino *et al.* 2015). Of the investigations which resolved Atlantic-Mediterranean partitioning, those that included fine scale sampling identified the Almeria-Oran front as the primary region driving this divergence in epipelagic species (Naciri *et al.* 1999; Perez-Losada *et al.* 1999, 2002; Quesada *et al.* 1995; Lemaire *et al.* 2005), whereas in deeper living species it was the shallow Gibraltar Strait that disrupts dispersal (Catarino *et al.* 2015). The final pattern of phylogeographic partitioning was where patterns of genetic differentiation did not correspond to a simple Mediterranean-Atlantic divergence, for example differentiation of eastern Mediterranean clades from a widespread western Mediterranean and Atlantic lineage (e.g. *Scomber scombrus* - Zardoya *et al.* 2004). Overall the wealth of investigation on diversification across the Atlantic-Mediterranean transition, and the contrasting evolutionary scenarios reported, highlight the limitation of making generalisations about a species response to a biogeographic boundary based on the patterns observed in closely related taxa, as historical changes and contemporary oceanographic features can have profoundly different demographic and evolutionary consequences on even closely related and co-distributed taxa (Bargelloni *et al.* 2003; 2005; Patarnello *et al.* 2007).

1.3.2 The Canary Current upwelling system

When compared with the other major upwelling currents globally (e.g. the Benguela, California and Humboldt currents) patterns of eco-evolutionary diversification across the CCUS remain remarkably poorly characterised. The majority of investigations that include the CCUS focus on much wider oceanic or inter-oceanic scales, with patterns of genetic partitioning across the CCUS assessed on comparisons of the north east Atlantic with a single sample from

tropical west Africa (Dillane *et al.* 2005; Catarino *et al.* 2015; Silva *et al.* 2017). Across these three studies different patterns of population genetic and phylogeographic structuring were reported, including: 1) genetic isolation of populations in the northeast Atlantic/ Mediterranean from those in tropical west Africa in the lesser flying squid (*Todaropsis eblanae* - Dillane *et al.* 2005). 2) genetic homogeneity across the entire Atlantic basin in the Portuguese dogfish (*Centroscymnus coelolepis* - Catarino *et al.* 2015); and 3) more complex patterns of historical diversification followed by recolonizations and homogenisation of geographic population differentiation in the European anchovy (*Engraulis encrasicolus* - Silva *et al.* 2017).

Most population genetic and phylogeographic investigations of the region at a smaller geographical scale have focussed on the northern Canary Current subsystem (north of the Cape Verde Front), and were directed at understanding patterns of diversification amongst the northern Macaronesian islands (i.e. the Azores, Madeira and the Canaries) and coastal eastern Atlantic and Mediterranean locations (Sá-Pinto *et al.* 2005; 2008; Domingues *et al.* 2007a;b; Zulliger *et al.* 2009; Calderon *et al.* 2008; Duran *et al.* 2004; Perez-Losada *et al.* 2007; Borrero-Perez *et al.* 2011). Most investigations of the northern Canary Current subsystem have described a lack of genetic partitioning amongst populations across this region (*Coryphoblennius galerita*: Domingues *et al.* 2007b; *Pollicipes pollicipes*: Quinteiro *et al.* 2007; *Megabalanus azoricus*: Quinteiro *et al.* 2015; *Helicolenus dactylopterus*: Aboim *et al.* 2005; *Phorbas fictitius*: Xavier *et al.* 2010), attributed to the high dispersal potential across coastal and island locations facilitated by the flow of the Azores Current and Canary Current, as well as the similar oceanographic conditions and geological history across the whole northern Canary Current subsystem.

Few studies have investigated patterns of phylogeographic and population genetic partitioning in taxa distributed across both the northern and southern Canary Current subsystems (defined here as the regions to the north and south of the Cape Verde front respectively), with the Cape Verde Islands constituting the single sample location from the southern Canary Current subsystem in the majority of studies that do (e.g. *Pollicipes pollicipes* - Quinteiro *et al.* 2007; *Megabalanus azoricus* - Quinteiro *et al.* 2015; *Helicolenus dactylopterus* - Aboim *et al.* 2005). These investigations all identified concordant isolation of the Cape Verde Islands from locations further north, dated to ~ 4.4MYA in the stalked barnacle *Pollicipes pollicipes* (Quinteiro *et al.* 2007), which was attributed to the persistent convergence zone of tropical and temperate water masses across the Cape Verde front and the different oceanographic regimes in the northern and southern Canary Current subsystems. However in a later study of the stalked barnacle, Campo *et al.* (2010) described a lack of genetic partitioning across

north-west African (Mauritania and Morocco) and Portuguese coasts, indicating that the isolation of Cape Verdean populations (at least in Quinteiro *et al.* 2007) is most likely attributed to particular geological and oceanographic features of this archipelago rather than the influence of the Cape Verde front. It seems therefore, that using the Cape Verde Island populations as a proxy for patterns of coastal diversification across the whole Canary Current system is dubious, and so the very limited data available point to a lack of genetic differentiation of marine populations along NW African and SW European coasts influenced by the oceanographic features of the Canary Current system.

Direct studies of phylogeographic and population genetic partitioning across the north west African coast at a local scale have been carried out only in a handful of species (*Pollicipes pollicipes* - Campo *et al.* 2010; *Sardina pilchardus* - Chlaida *et al.* 2006, Gonzalez & Zardoya 2007, Atarhouch *et al.* 2007; *Mytilus galloprovincialis* - Jaziri & Benazzou 2002). In the European sardine (*Sardina pilchardus*) allozyme and EPIC-PCR (Emulsion, Paired Isolation and Concatenation PCR) investigations identified a genetic break along the coast of Morocco, associated with the Ghir upwelling cell, with sardine populations subdivided into north east Atlantic and southern Atlantic (southern Morocco to Mauritania) groups (Atarhouch *et al.* 2006, 2007; Chlaida *et al.* 2006 - but see Gonzales & Zardoya 2007). A break in connectivity in association with the Ghir upwelling was also identified in *Mytilus galloprovincialis* populations across the Moroccan coastline (Jaziri & Benazzou 2002).

It would therefore appear that for some species that inhabit the CCUS, the upwelling cells distributed across the north west African coastline and/or the Cape Verde Front represent a significant barrier to historical and contemporary dispersal, promoting genetic partitioning (Chlaida *et al.* 2006; Gonzalez & Zardoya 2007; Atarhouch *et al.* 2007; Jaziri & Benazzou 2002; Quinteiro *et al.* 2007; Aboim *et al.* 2005). However for other species genetic homogeneity was reported across the entire CCUS (Campo *et al.* 2010; Catarino *et al.* 2015; Silva *et al.* 2017). Regardless of the patterns of contemporary population differentiation reported across the CCUS it appears that this area and further south into tropical west Africa has represented an important refuge for northeast Atlantic and Mediterranean taxa during Pleistocene glacial maxima. The northeast Atlantic and Mediterranean were profoundly, and often directly, influenced by glacial advancement (Svendsen *et al.* 2004; Section: 1. 2), with genetic signatures of localised extinctions and latitudinal range shifts reported for numerous taxa across these marine provinces compared to higher diversity and diversification retained within the areas south of the Cape Verde Front (e.g. Almada *et al.* 2001; Schiebel *et al.* 2002; Domingues *et al.* 2006; 2007a; Quinteiro *et al.* 2007; 2015; Campo *et al.* 2010).

1.3.3 Tropical west Africa

Unlike the other east Atlantic Ocean regions examined in this thesis, the tropical coast of west Africa (defined here as Senegal to Angola) experienced limited direct impact of Pleistocene glacial periods on its physical environment (see section 1.2). However, the heterogeneous oceanographic features of this region (section 1.3) in combination with the proposed range shifts of higher latitude species into tropical waters during glacial maxima (Domingues *et al.* 2006; 2007a; Quinteiro *et al.* 2007; 2015; Campo *et al.* 2010) make this an interesting region to examine the evolutionary history of taxa distributed across the tropics and around the northern and southern periphery of this province.

For temperate species residing at the fringes of the tropical eastern Atlantic province, periods of global cooling likely drove trans-equatorial migrations, giving rise to the anti-tropical distributions observed in numerous eastern Atlantic taxa. Although the processes shaping these anti-tropical distributions have been well characterised for Pacific taxa (Burridge 2002; Ludt *et al.* 2015; Ludt 2018), no molecular genetic investigations of anti-tropicality have been conducted within the eastern Atlantic. However, as with the Pacific system these distributions are likely explained by two predominant mechanisms: 1) dispersal of a species across the tropics, facilitated by either isothermal submergence (e.g. Hubbs 1952; Ekman 1953), human mediated transport (e.g. Carlton & Geller 1993) or reduced tropical SST associated with periods of cooler global climate (e.g. Ekman 1953; Lindberg 1991); or 2) vicariance-driven expulsion of tropical taxa to higher latitudes, driven by either the mid-Miocene rise in tropical temperatures (e.g. Valentine 1984; White 1986) or competitive exclusion by a newly emerged tropical species (e.g. Briggs 1987). Although trans-equatorial dispersal during periods of global cooling appears to be the most common driver of anti-tropicality in Pacific taxa (Burridge *et al.* 2002; Ludt *et al.* 2015), without comprehensive phylogenetic and phylogeographic characterisation of eastern Atlantic taxa, with corresponding molecular dating that can be reliably attributed to major climatic/ geological events, the historical and contemporary permeability of the west African tropics to temperate Atlantic species is unknown.

For those species that are adapted to, and whose distributions extend throughout, the tropical eastern Atlantic, patterns of eco-evolutionary diversification are more likely shaped by the oceanographic conditions intrinsic to this region and their interaction with historical climatic and geological events. The marine fauna of tropical coasts of west Africa is poorly characterised in regards to the processes shaping patterns of eco-evolutionary diversification, particularly when compared with other tropical regions globally. However, with the few molecular investigations of its taxa to date (Falk *et al.* 2003; Durand *et al.* 2005b; 2013; Reid *et al.* 2016;

Silva *et al.* 2017) some commonalities in patterns of phylogeographic diversification across west Africa, and their interaction with species life history and ecology, can be inferred.

For small pelagic fish species with broad Atlantic or interoceanic distributions the tropical coast of west Africa does not appear to represent a significant biogeographic barrier, with historically stable and panmictic populations reported across the entire eastern Atlantic (e.g. Silva *et al.* 2017), most likely resulting from the large effective population sizes of these species. However for larger pelagic species such as the bluefish (*Pomotomous salatrix)* subtle phylogenetic partitioning has been reported across the tropical west African coast with no clear geographical pattern, which was attributed to historical population fragmentation, diversification and subsequent secondary contact (Reid *et al.* 2016). For the bluefish, as has been reported in other coastal and anadromous fish species (Falk *et al.* 2003; Durand *et al.* 2005b, 2013), it is likely that range shifts and localised extinctions during Pleistocene glacial and interglacial cycles have driven the reported patterns of phylogeographic partitioning. Refuge areas for tropical taxa have been proposed in several areas along the west African coast, including Angola, Cameroon/Gabon, eastern Ivory Coast/western Ghana, and Sierra Leone/ Liberia (Falk *et al.* 2003).

Investigations of fine scale patterns of population connectivity along the west African coast have focussed primarily on estuarine-associated fish species (e.g. *Sarotherodon spp. complex - Falk et al. 2003; Ethmalosa fimbriata* - Gourene *et al.* 1993 and Durand *et al.* 2005b, 2013*)*, with generally congruent patterns of intra-specific partitioning reported. Isolation of populations within the Gulf of Guinea from those further north (Ivory Coast to Mauritania) and south (Congo to Angola) was resolved in all estuarine associated species (Gourene *et al.* 1993; Durand *et al.* 2005b; 2013; Falk *et al.* 2003), with the isolation of northern populations from the Gulf of Guinea also reported in the small pelagic round sardine (*Sardinella aurita* - Chikhi *et al.* 1998). The clear concordance in the response of tropical eastern Atlantic fish species to the historical and contemporary environment of the west African coast indicates that although this region remains poorly investigated, its biota have been subjected to concurrent, persistent and universal drivers of phylogeographic diversification that remain to be appropriately characterised.

1.3.4 The Benguela Upwelling System & Atlantic-Indian Ocean boundary

The southernmost tip (Cape Agulhas) of the coast of South Africa forms the solitary point of direct contact between the Atlantic and Indian Oceans, and as the waters at this latitude are warm temperate this region potentially allows contact between the warm temperate and subtropical fauna of the two oceans. However dispersal between these ocean basins appears to be limited for numerous marine species, resulting in the widely recognised South African

marine biogeographic breakpoint (Turpie *et al.* 2000; Lessios *et al.* 2003; Floeter *et al.* 2008). Studies employing wide geographic (often global, inter-oceanic) sampling have demonstrated a breakdown in gene flow between Indian Ocean and Atlantic populations, often accompanied by speciation events, that can be associated with the convergence of cool upwelled water from the BUS with the tropical Indian Ocean waters of the Agulhas Current (Lessios *et al.* 2003; Barth *et al.* 2017; Mullins *et al.* 2018).

Despite the recognition of the Benguela Current as one of the most important biogeographic barriers in the Atlantic Ocean, few studies have incorporated the entire Benguela region, with the majority of population genetic and phylogeographic research conducted across the southern Benguela subsystem alone (Harrison 2002; Evans *et al.* 2004; Teske *et al.* 2007a, b; Zardi *et al.* 2007; Von der Heyden *et al.* 2008). The northern and central areas of the Benguela system, extending from Angola to western South Africa on the other hand have received comparatively limited attention in regards to patterns of eco-evolutionary diversification. Therefore the historical and contemporary impact of the cool Benguela Current and warm boundary Angola and Agulhas Currents on the demography and diversification of marine species across this whole region remains relatively unexplored.

Those investigations that have included samples that span the whole BUS (e.g. Henriques 2012; Henriques *et al.* 2012, 2014, 2016; DeBeer 2014; Reid *et al.* 2016; Soekoe 2016;Gwilliam 2017; Gwilliam *et al.* 2018) have found generally concordant patterns of phylogeographic and population genetic diversification, that can be associated directly with the main perennial upwelling region found along the coast of central and southern Namibia. Furthermore, for numerous taxa it would appear that prolonged isolation of populations to the north and south of the upwelling region has resulted in speciation events, or at least substantial intraspecific divergence. Where dated, intra-specific subpopulation isolation across the BUS occurred commonly within the mid to late Pleistocene between 200 and 700KYA, a period of profound vicariance within the BUS (see section 1.2), for example in fish species: *Diplodus capensis* (Henriques 2012), *Lichia amia* (Henriques *et al.* 2012), *Pomatomus saltatrix* (Reid *et al.* 2016), *Lithognathus mormyrus* and *Sarpa salpa* (Gwilliam 2017), and the cephalopod *Octopus vulgaris* (DeBeer 2014). Species-level diversification events, however, were associated with the late Pliocene / early Pleistocene approximately 2-3 MYA, coinciding with the establishment of the contemporary major oceanographic features of the BUS region (~2MYA), for example in fish species like *Atractoscion* spp (Henriques *et al.* 2014), *Argyrosomus* spp (Henriques 2012) and *Spondyliosoma* spp. (Gwilliam 2017).

Despite the evidence for long-term isolation of populations to either side of the BUS in repeated examples tested, the BUS might not have been an entirely impermeable and persitent

barrier to dispersal since the formation of its contemporary features. Genetic signatures of asymmetric dispersal from South African populations to those to the north of the BUS, and subsequent secondary contact and interbreeding, have been reported in *Pomatomous saltatrix* (Reid *et al.* 2016), *Thunnus albacares* (Barth *et al.* 2017; Mullins *et al.* 2018), *Lithognathus mormyrus* and *Sarpa salpa* (Gwilliam 2017). It would also appear that for many populations to the north and south of the BUS, whether isolated or not, the termination of Pleistocene glacial-interglacial cycles promoted substantial demographic change, with signals of population expansions after the Last Glacial Maximum (likely following a LGM-associated bottleneck) reported for many coastal and pelagic fish species (Henriques 2012; Henriques *et al.* 2012, 2014, 2018; DeBeer 2014; Gwilliam 2017). Although patterns of population genetic diversification across the BUS indicate that this oceanographic feature has been and is a major phylogeographic boundary, the species that inhabit this region exhibit an array of responses to the historical variability of this environment and its contemporary oceanography, raising questions regarding the commonality in timings of diversification or demographic fluctuations and their association with historical climatic/geological events.

1.4 Aims and Objectives

Regardless of whether the molecular genetic basis of diversification has been well characterised across a region, the presence of multiple taxa with shared distributional limits indicates the presence of features or processes intrinsic to that region that have broadly shaped the evolutionary history of its fauna. As such the biogeographic partitioning of the eastern Atlantic (based on levels of endemism and species occurrence - Spalding *et al.* 2007) into four main provinces (Lusitanian, Gulf of Guinea, Benguela and Agulhas) indicates common features at the intersection of these provinces that serve to disrupt species distributions as well as historical and contemporary migration and consequent gene flow. Furthermore the occurrence of parallel distributions of multiple sister species and species complexes between the temperate waters of the Northeast Atlantic, through the tropics of west Africa and into South Africa have been observed across an array of taxa, for example in crustaceans such as krill (D'Amato *et al.* 2008) and spiny lobster (Palero *et al.* 2009), molluscs such as octopus (Teske *et al.* 2007b), and fish such as anchovy (Grant & Bowen 2006) and hake (Grant & Leslie 2001). The similar phylogeographic patterns observed across these species pairs and complexes, which generally form a Mediterranean-northeast Atlantic group or clade, a tropical west African group/clade and a South African group/clade, further support the proposed presence of shared evolutionary pressures faced by eastern Atlantic taxa and commonalities in the drivers of diversification and speciation along the west African coast.

Through assessing shared patterns in the distribution of intra- and inter-specific genetic diversity across multiple co-distributed taxa, common features and processes driving the divergence of species, alongside the temporal basis of these divergence events can be identified (Avise 2000; Bermingham & Moritz 1998; Bowen *et al.* 2016). Accordingly the present study set out to assess shared patterns of phylogeographic and population genetic partitioning in a range of species across the eastern Atlantic, a historically fluctuating environment in regards to its geology, climate and oceanography. Through the investigation and comparison of multiple widespread species and species complexes with shared distributions but contrasting life history characteristics, the aim was to assess the relationship between the oceanographic features and climatic history of the eastern Atlantic and the processes driving genetic divergence and reproductive isolation in taxa across this region. As one of the most persistent and commonly recognised phylogeographic boundaries in the eastern Atlantic, the BUS is a primary focus of the study. But the aim was to choose study species /species complexes with distributions extending from the Northeast Atlantic and Mediterranean to the Southwest Indian Ocean coasts of South Africa in order to assess the interaction of the other major oceanographic boundaries across the eastern Atlantic (and effects of past climatic changes on these areas) with processes of genetic differentiation and speciation occurring across the Benguela System. The study taxa selected (based on their contrasting life history characteristics and ecologies) were: i) the nektobenthic and resident Common Cuttlefish (*Sepia officinalis*) species complex (Chapters 2.1, 2.2, 2.3); ii) the pelagic, catadromous grey mullets (*Mugilidae*) (Chapter 3.1 & 3.2); iii) the pelagic and highly migratory horse mackerel *(Trachurus trachurus*) species complex (Chapter 4); and iiii) the demersal *Dentex dentex* species complex (Chapter 6).

Furthermore through comparing the resolved phylogeographic structure of taxa that reside within the eastern Atlantic (a historically fluctuating environment), with those resolved for widespread coastal fish species distributed across the southwest Indian Ocean (*Lethrinus harak*, *Lethrinus nebulosus* and *Lethrinus mahsena* - Chapters 5.1 & 5.2), an additional objective was to assess phylogeographic structuring and historical demographic processes in a speciose and widespread coastal fish group (Lethrinids) across the southwest Indian Ocean, a comparatively more stable environment historically (Ridgway & Sampayo 2005). The aim was to produce a baseline scenario that facilitates comparisons regarding the true extent to which past geological and climatic events influenced speciation and divergence processes in the eastern Atlantic.

Although a number of studies have investigated the influence of a single proposed eastern Atlantic biogeographic boundary on the phylogeographic and population genetic partitioning

of taxa across this region (see section 1.3), few have examined the region as a whole. Therefore through the inclusion of samples from across the Mediterranean-northeast Atlantic, Canary Current Upwelling System, tropical west Africa and the Benguela Upwelling System the aim is to understand the inter-relationship (if any) of these various proposed phylogeographic / biogeographic break-points and the order and direction of speciation/divergence events across the eastern Atlantic coast of Africa. By comparing the present study results with those of previous studies, an overarching aim was to elucidate the relative impact of these proposed biogeographic boundaries and historical climatic events on the demographic and evolutionary history of eastern Atlantic coastal marine taxa. As the Benguela System and the coasts of SW Africa currently represent the main point of potential connection and interaction between the marine fauna of the Atlantic and Indo-Pacific Oceans, the present study may provide some information on how species may cross and/or evolve in isolation around this globally important biogeographic boundary. All of the study species represent locally important fishery resources, so an additional interest was to investigate how fragmentation of species distributions and habitats across the eastern Atlantic may influence the vulnerability of genetically isolated species and populations to future climatic changes and fishing pressures.

CHAPTER 2

Evolutionary history of the *Sepia officinalis species complex across the eastern Atlantic and Mediterranean*

Published in: African Journal of Marine Science (2017)
DOI: 10.2989/1814232X.2017.1371076

CHAPTER 2.1

Phylogeny of the *Sepia officinalis* species complex in the eastern Atlantic extends the known distribution of *Sepia vermiculata* across the Benguela upwelling region.

2.1.1 Introduction

As many traditionally exploited fin fish stocks continue to decline there is growing interest in the expansion of cephalopod fisheries (Boyle 2000, Young *et al.* 2006, Anderson *et al.* 2011, Jereb *et al.* 2015). The typical short life cycle of cephalopods renders them vulnerable to overfishing (Rodhouse *et al.* 2014) and as they fulfil important roles in marine ecosystems, improved assessment and management of stocks will be vital to ensure ecosystem compatible exploitation (Pierce *et al.* 1998, Young *et al.* 2006). Fundamental to this is both accurate species identification and resolution of species ranges (Taylor *et al*. 2012, McKeown *et al*. 2015).

The common cuttlefish (*Sepia officinalis*) species complex is of importance to both commercial and artisanal fisheries across its range (Reid *et al.* 2005). Three species are currently described within this complex (Khromov *et al.* 1998): *Sepia officinalis* (Linneaeus 1758), *Sepia hierredda* (Rang 1837) and *Sepia vermiculata* (Quoy & Gaimard 1832). By far the most extensively studied of these species is *S. officinalis*, an abundant cephalopod within coastal waters of the Mediterranean Sea basin and north east Atlantic Ocean. The northern distribution of *S. officinalis* extends into the southern North Sea (Gittenberger & Schrieken 2004, De Heij & Baayen 2005), and the southern limits are along the north west coast of Africa, coinciding with the border between Mauritania and Senegal (16°N). Off North West Africa *S. officinalis* is found in sympatry with *S. hierredda*, the distribution of which extends as far north as Cape Blanc (21° N)(Hatanaka 1979, Guerra *et al.* 2001). *S. hierredda* is found at greater depths than *S. officinalis* and although it is relatively well characterised in its zone of overlap with *S. officinalis* (e.g. Guerra *et al.* 2001) there has been limited, if any, research focussed upon *S. hierredda* from its central or southern distribution. Despite this, fisheries data cite the distribution of *S. hierredda* as extending throughout the tropics and subtropics as far south as Tigres Bay in southern Angola (Hatanaka 1979, Roeleveld 1998). A break in the occurrence of this species complex is noted around the Benguela Upwelling System (BUS) that occurs off the coast of Namibia, with *S. vermiculata*, the most poorly investigated member of this species complex, thought to be restricted to the coast of southern Africa, occurring from the Western Cape of South Africa into the Indian Ocean as far as central Mozambique (Roeleveld 1972, 1998, Khromov 1998). Additionally trawl data from far further into the Indian Ocean

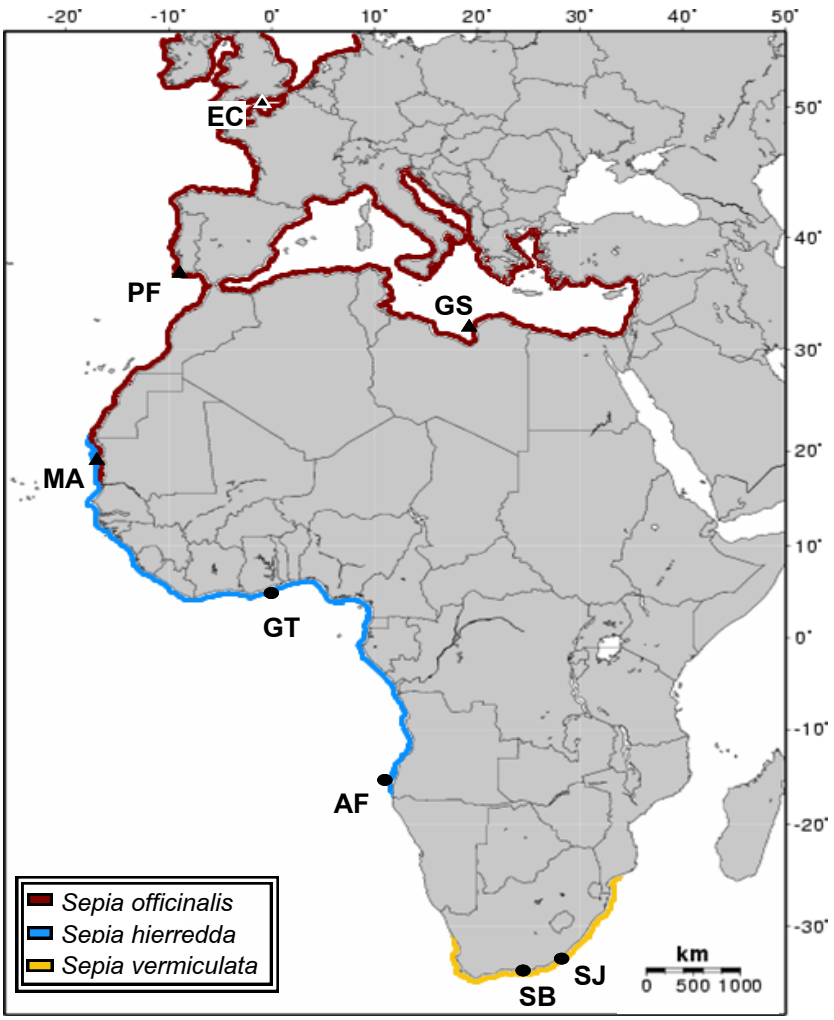


Figure 2.1.1: Sampling sites for *S. vermiculata* and *S. hierredda* across the south-east Atlantic and Indian Oceans (denoted by circles; GT=Tema, Ghana; AF=Flamingo, Angola; SB=Bushmans river, South Africa; SJ=Jefferys Bay, South Africa), as well as locations of north-west Atlantic and Mediterranean sequences of *S. officinalis* taken from GENBANK (denoted by triangles; MA=Mauritania; PF=Faro, Portugal; EC=English Channel; GS=Gulf of Sidra). Coloured areas represent the currently recognised distribution of each species.

noted the occurrence of a population of *S. vermiculata* on the Saya-de-Malha Bank of the Mascarene plateau (Nesis 1993).

While available genetic data support the distinctiveness of *S. officinalis* and *S. hierredda* (Guerra *et al.* 2001), at present the description of *S. vermiculata* is based solely on divergence from *S. hierredda* and *S. officinalis* in morphological traits (Khromov *et al.* 1998). As such a primary goal of the present study was to assess the validity of *S. vermiculata* as a species using mitochondrial DNA (mtDNA) sequencing. A secondary objective was to assess genetic patterns in the context of biogeography, as to date and to the best of our knowledge, there has been no molecular investigation of the *S. officinalis* species complex south of Mauritania. The results, based on mtDNA COI and cytb sequencing, support the species status of *S. vermiculata* but indicate that its range extends further north in the Atlantic Ocean than previously described, and at least as far as southern Angola.

2.1.2 Methods

2.1.2.1 Sampling and mtDNA sequencing

Tissue samples (tentacle clips stored in 95% ethanol) recorded as *S. hierredda* were collected between 2011 and 2016 from artisanal catches in Ghana (Tema fish market) as well as through targeted fishing in southern Angola (Flamingo), while tissue samples recorded as *S. vermiculata* were obtained from two locations (Bushmans River and Jeffreys Bay) in the Eastern Cape of South Africa (Fig. 2.1.1).

Genomic DNA was extracted from all samples using a standard CTAB-chloroform/ isoamylalcohol method (Winnepenninckx *et al.* 1993). Partial sequences of the mtDNA cytochrome c oxidase subunit 1 (COI) and the Cytochrome oxidase B (cytb) genes were amplified by polymerase chain reaction (PCR) using species-specific primers developed specifically for this study (COI: SepiaCOIF 5'-GTAAACCTGGTACACTTTT-'3, SepiaCOIR 5'-TTCTATTTGTAAACCTTCTCATC-'3; Cytb: cytb117F 5'-CCCCCAATCCAAGTTAACAA-'3, cytb928R 5'-ATGCGGGATGTGAATTATGG-'3). PCRs were performed in a total volume of 20µl, containing 4 µL template DNA, 2 mM $MgCl_2$, 0.5 µM forward primer and 0.5 µM of reverse primer, 0.2 mM dNTP mix (20 µM each dATP, dCTP, dGTP, dTTP), 1x reaction buffer [75 mM Tris-HCl, 20 mM $(NH_4)_2SO_4$] and *Taq* polymerase (BIOTAQ, 5 U/µl). The PCR thermo-profile for COI amplification was: 180s at 95°C, followed by 40 cycles of 30s at 95°C, 45 seconds annealing at 50°C and 60s at 72°C, followed by a final 5 minute extension at 72°C. For cytb amplification PCR conditions were: 180s at 95°C, followed by 34 cycles of 30s denaturing at 95°C, 30s annealing at 52°C and 30s at 72°C, again followed by a final 5 minute extension at 72°C. PCR products were then purified and sequenced using BigDye technology, with sequence identity confirmed using BLAST.

Table 2.1.1: Collection locality and GENBANK accession numbers (where applicable) for all samples of the *Sepia officinalis* species complex used in this investigation, * denotes where sequences were obtained from GENBANK.

country	location	Code	N (COI)	N (cytB)	GenBank accession numbers
South Africa	Jeffreys bay	SJ	8	7	
South Africa	Bushmans river	SB	10	8	
Angola	Flamingo	AF	10	6	
Ghana	Tema	GT	8	8	
Mauritania		MA	4*		EF416525-Ef416528
Portugal	Faro	PF	4*		EF416384-EF416387
English channel		EC	4 *	1	EF416306-EF416309
Gulf of Sidra		GS	4*		EF416535-EF416538
unknown		S. officinalis		2*	AB240155, NC007895
unknown		S. pharaonis	1*	1*	NC02146

2.1.2.2 Phylogenetic sequence analysis

Phylogenetic relationships among sequences obtained here, as well as other sequences available on GENBANK (Table 2.1.1) were inferred using maximum likelihood (ML) trees, constructed for both mtDNA regions in MEGA v6.06 (Tamura *et al.* 2013) and Bayesian inference performed using MRBayes v3.2 (Ronquist & Huelsenbeck 2003). In both cases HKY+G+I was identified as the best fit substitution model based on the Akaike Information Criterion (AIC; Akaike 1974) implemented in MODELTEST. For both gene regions *Sepia pharaonis* was used as an outgroup as it was the most closely related species for which COI and cytb sequences were available. Maximum likelihood bootstrap values were calculated using 1000 bootstrap replicates and Bayesian Inference (BI) was calculated assuming unknown model parameters, and run over 5,000,000 generations, sampling the Markov chain every 1000 generations and using 3 heated chains and one cold chain. It was considered that convergence had been reached on the basis that the standard deviation of split frequencies was <0.01, with the first 15% of trees discarded as burn-in. Percentage sequence divergence within and between species/clades were calculated using MEGA v6.06.

2.1.3 Results

In total 52 individuals were sequenced for COI (345bp) and 32 individuals sequenced for cytb (500bp). Phylogenetic analysis of all sequences revealed 3 strongly supported clades for both mtDNA regions, corresponding to the three described species of *S. officinalis, S. hierredda* and *S. vermiculata* (Fig. 2.1.2 and Fig. 2.1.3). COI and cytb sequences of 8 individuals from Ghana yielded 2 and 6 haplotypes respectively, which aligned with *S. hierredda* according to BLAST searches. All COI sequences from South Africa (N=18) and Angola (N=10) yielded a single haplotype, and based on phylogenetic placement were concluded to be *S. vermiculata* (Fig. 2.1.2). For the cytb sequences of 15 individuals from South Africa 2 haplotypes were present, with an additional 4 haplotypes resolved within the cytb data set for the 6 individuals sequenced from Angola: again all Angolan haplotypes clustered with the South African *S. vermiculata* haplotypes (Fig. 2.1.3).

Sequences that fell within the *Sepia officinalis* clade were from locations north of Mauritania, including the English Channel and Mediterranean. As Perez-Losada *et al.* (2007) demonstrated in their original analysis of the COI sequences used here, high levels of intra-specific phylogenetic structuring was observed within *S. officinalis*, with 3 well-supported COI clades (BI=0.82-0.89, BS=86-99) observed in the subset of COI sequences used in the present analysis. However within *S. hierredda* and *S. vermiculata* low levels of phylogenetic diversification were ob-

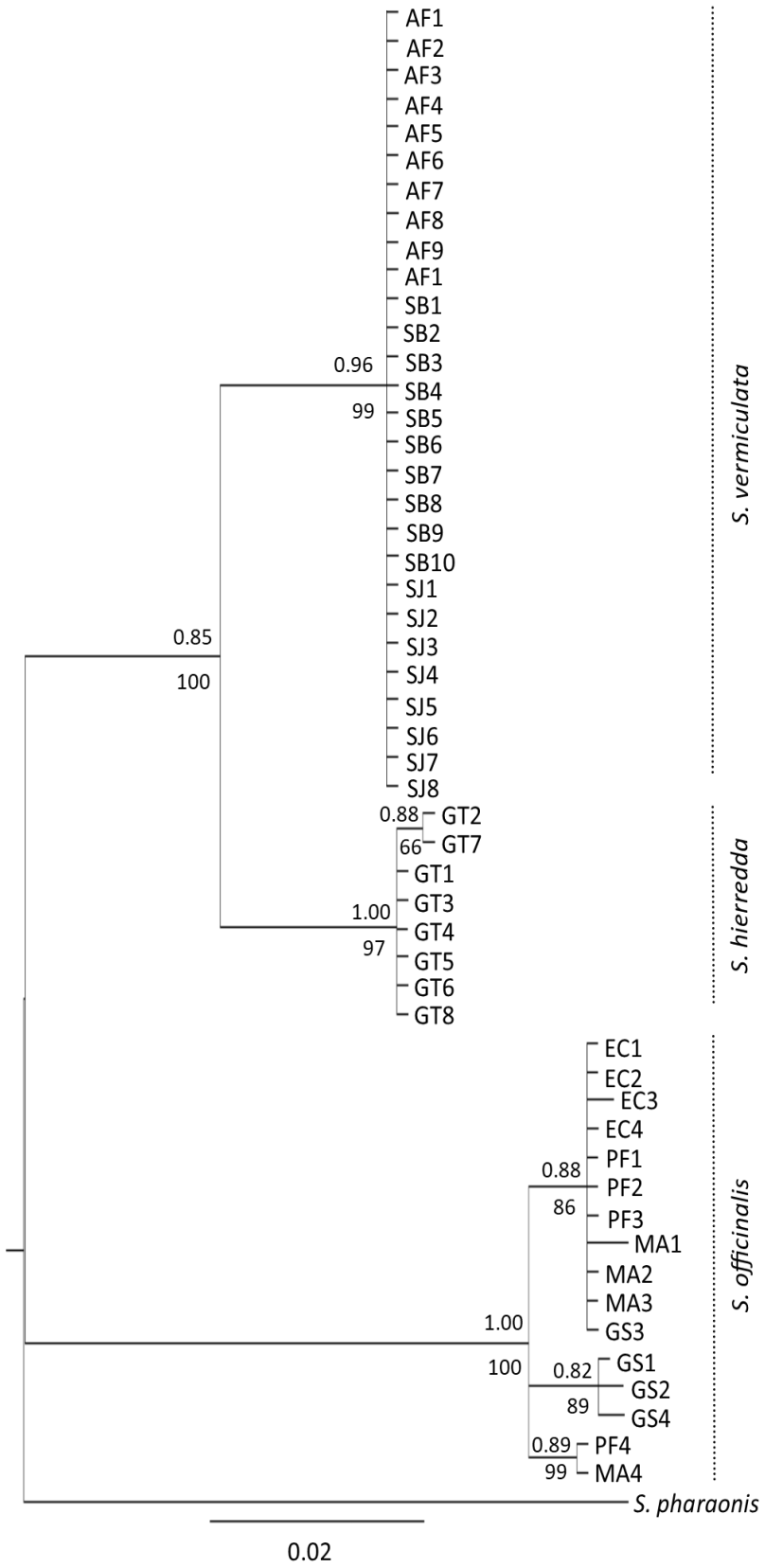


Figure 2.1.2: Bayesian phylogram depicting the relationships between *Sepia officinalis*, *Sepia hierredda* and *Sepia vermiculata* sampled across the east Atlantic, Mediterranean and Indian Oceans based upon partial sequences of the mtDNA COI gene. Bayesian Inference posterior probabilities are shown above nodes and maximum likelihood bootstrap values are given below. Branch lengths are proportional to the number of nucleotide substitutions and *Sepia pharaonic* is included as an out-group species. Taxon codes refer to locations in Figure 1.

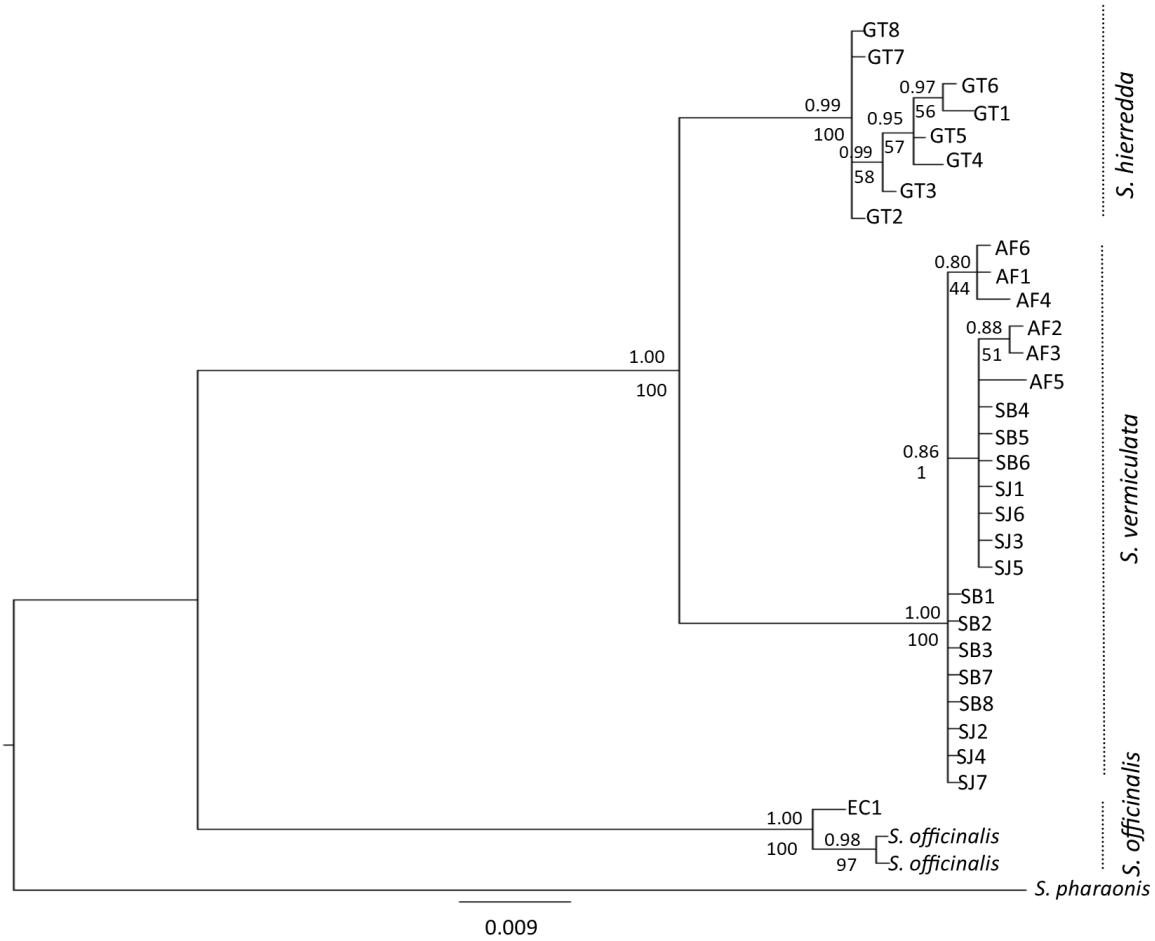


Figure 2.1.3: Bayesian phylogram depicting the relationships between *Sepia officinalis*, *Sepia hierredda* and *Sepia vermiculata* sampled across the east Atlantic, Mediterranean and Indian Oceans based upon partial sequences of the mtDNA cytb gene. Bayesian Inference posterior probabilities are shown above nodes and maximum likelihood bootstrap values are given below. Branch lengths are proportional to the number of nucleotide substitutions and *Sepia pharaonis* is included as an outgroup species. Taxon codes refer to locations in Figure 1.

served using COI. The cytb dataset was comparatively more variable than that of COI, with greater levels of intra-specific genetic divergence observed. This was particularly obvious in *S. vermiculata* where the Angolan sample (a single COI haplotype in Angolan and South African samples) comprised 4 private haplotypes with moderate support for the divergence of this Angolan sample from the South African sample (BI=0.80-0.88, BS=44-51).

Inter-specific genetic distances (percentage sequence divergence) were greatest between *S. officinalis* and *S. vermiculata* in both the COI (Table 2.1.2A) and cytb (Table 2.1.2B) datasets (COI=13.37%, cytb=12.20%), followed by *S. officinalis* and *S. hierredda* (COI=11.37%, cytb=11.71%), with *S. hierredda* and *S. vermiculata* the least genetically different (COI=5.72%, cytb=4.83%). Comparatively intraspecific genetic distances were low for all three species, ranging from 0-1.12% for COI and 0.24-0.53% for cytb.

Table 2.1.2: Pairwise genetic distances between *Sepia officinalis*, *Sepia hierredda* and *Sepia vermiculata* based on partial sequences of A) the mtDNA COI gene, and B) the mtDNA cytb gene. Percentage sequence divergence between putative species/clades are given below the diagonal with *P*-distances above the diagonal. Intraspecific percentage sequence divergence is on the diagonal. Standard error for all distance values are given in parentheses

A)	*S.officinalis*	*S.hierredda*	*S.vermiculata*
S. officinalis	0.012 (0.003)	0.114 (0.015)	0.134 (0.017)
S. hierredda	0.139 (0.023)	0.001 (0.001)	0.057 (0.012)
S.vermiculata	0.172 (0.029)	0.064 (0.015)	0.000 (0.000)

B)	*S.officinalis*	*S.hierredda*	*S.vermiculata*
S. officinalis	0.005(0.003)	0.117 (0.013)	0.122 (0.014)
S. hierredda	0.153 (0.024)	0.004 (0.002)	0.048 (0.009)
S.vermiculata	0.166 (0.027)	0.054 (0.012)	0.002 (0.001)

2.1.4 Discussion

Phylogenetic analysis of two mtDNA genes resolved three highly supported and reciprocally monophyletic clades corresponding to *S. officinalis*, *S. hierredda* and *S. vermiculata*. Applying phylogenetic species criteria, this result represents the first molecular genetic confirmation of the distinct species status of *S. vermiculata*. This conclusion was further supported by inter-specific genetic distances which were greater than those observed between other closely related but taxonomically distinct cephalopod species (Dai *et al*. 2012, Amor *et al.* 2015), as well as ratios of within- to between-species DNA sequence divergence which were in excess of commonly applied species barcoding ratios (Hebert *et al.* 2004, Meyer & Paulay 2005, Lefebure *et al.* 2006).

Interestingly, and of pertinence to fisheries management of these species, the data presented here show that the distribution of *S. vermiculata* extends further north than previously described, with all samples from southern Angola falling within the *S. vermiculata* clade in both the COI and cytb datasets. Prior to this investigation *S. vermiculata* was considered to be a South African (and Indian Ocean) endemic whose extension northward along the west African coast appeared to be limited to southern Namibia by the cold waters of the BUS (Roeleveld 1972, 1998). However the coastal areas of Angola have received comparatively limited prior research, particularly in relation to the abundance and distribution of cephalopods, with the only mention of Angolan cuttlefish coming from the bottom trawl data of Bianchi (1992), where all *Sepia* caught were broadly classified as belonging to the *S. officinalis* species com-

plex. Therefore, it is possible that *S. vermiculata* caught in Angolan waters have been previously misidentified as *S. hierredda*. However as only 10 Angolan samples were included in the COI analysis, the absence of *S. hierredda* could also reflect a greater abundance of *S. vermiculata* and/or temporal variance in distribution coinciding with sampling sites. The misidentification of morphologically similar species and over/under representation of species richness and abundance can cause inaccuracies in our understanding of biological, ecological and evolutionary processes (Garcia-Vazquez *et al.* 2012, Tillett *et al.* 2012). Consequently a comprehensive genetic analysis of further spatial and/or temporal samples will be needed to accurately assess the extent of overlap or geographical separation between these cuttlefish species.

Despite an overall lack of genetic diversity and structuring in the COI data set of *S. vermiculata*, analysis of cytb sequences revealed some evidence of phylogenetic diversification between individuals from South Africa and Angola, which can be readily aligned with the oceanography of this region. The expanse of coastal habitat between South Africa and southern Angola is dominated by the Benguela Cold Current and the associated perennial upwelling system. The BUS is an area which owing to its persistent cool upwelled waters is generally considered to represent a biogeographic and evolutionary boundary region to many marine species (e.g. Henriques *et al.* 2014, 2016). More recently Reid *et al.* (2016) reported asymmetric gene flow across the BUS from South Africa into Angolan waters, indicating some degree of historical permeability to this system that may help explain the patterns observed here for *S. vermiculata*. This commonly observed restriction to gene flow in association with the BUS is likely enhanced in *S. vermiculata* by its life history characteristics, namely the lack of a highly dispersive pelagic larval stage (Perez-Losada *et al.* 1999, 2002, 2007, Boyle 2000). In order to comprehensively determine whether there is bi-parentally restricted gene flow across the BUS and indeed between the putative species designations of this study, analysis of nuclear genetic polymorphisms would be required. These findings thus highlight the need for a comprehensive phylogeographic and population genetic evaluation of *Sepia* across the southern African coast in order to fully characterise patterns of genetic connectivity and the drivers behind them.

Here we not only provide the first molecular confirmation of the species status of *S. vermiculata* but also extend this species' known geographical range within the east Atlantic from the west coast of South Africa (Roeleveld 1972, 1998, Reid *et al.* 2005) to southern Angola, and in doing so highlight the likely incidence of harvesting of misidentified species. This has implications for the management of *Sepia* in southern African waters, which will require a thor-

ough investigation of the abundance and distributional limits of both *S. vermiculata* and *S. hierredda* in order to appropriately conserve the biodiversity of this region and negate the detrimental impacts of indiscriminate harvesting. Finally we reveal subtle patterns of phylogenetic diversification between *S. vermiculata* from South Africa and Angola, indicating that as for many marine teleosts (Henriques *et al.* 2014, 2016) the BUS constitutes a biogeographic barrier to dispersal for the Sepiidae. Ultimately this investigation highlights the need for a thorough molecular examination of *Sepia* in west African waters and for this to be integrated into fisheries stock assessment, with the aim of not only determining the stock status of cuttlefish fisheries but also ascertaining the drivers that have promoted both inter- and intraspecific divergence within this species complex.

CHAPTER 2.2

Comparison of sequence variation among the duplicated non-coding regions of the cuttlefish mitochondrial genome

2.2.1 Introduction

Mitochondrial DNA (mtDNA) is a popular marker in phylogenetics, phylogeography and population genetics, owing to qualities, such as high mutation rate and ease of PCR amplification and sequence analysis (Avise 2000). The control region (CR)/ Non-coding region (NCR) harbours the regulatory elements required for the replication and expression of the mitochondrial genome (Shadel & Clayton 1997).However the CR/NCR does not code for a functional gene, and as such is generally expected to accumulate mutations more rapidly than other coding parts of the mtDNA genome. Consequently it offers the potential for enhanced resolution of genetic structuring over evolutionary (1000+ generations) and ecological (1-20 generations) time scales.

The length and structure of the CR/NCR is highly variable across taxa with length ranging from hundreds to tens of thousands of base pairs (Taylor *et al.* 1993; Kitazaki & Kubo 2010). In comparison with other taxa little is understood regarding the structure and function of the CR/NCR in marine invertebrates. The majority of investigations of CR/NCR divergence among marine invertebrates have focussed on species of decapod crustaceans (Chu *et al.* 2003; McMillen-Jackson and Bert, 2003, 2004; Diniz *et al.* 2005), and only a handful of investigations have assessed NCR variation in cephalopods (Aoki *et al.* 2008; Winkelmann *et al.* 2013). Such studies have reported striking cases of duplicated NCRs. The frequency of these NCRs across the mitochondrial genome appears to be genus/ family specific with a single NCR observed in Octopus, Vampyroteuthis and Nautilus species, two identified in Sepia, Sepioteuthis, Watesenia, Todarodes and Architeuthis species and three in Loligo species (Akasaki *et al.* 2006; Tomita *et al.* 2002; Yokobori *et al.* 2004; 2007; Winkelmann *et al.* 2013). Unlike other marine invertebrates, the NCR of cephalopods appears to exhibit a more complex pattern of evolution, with several

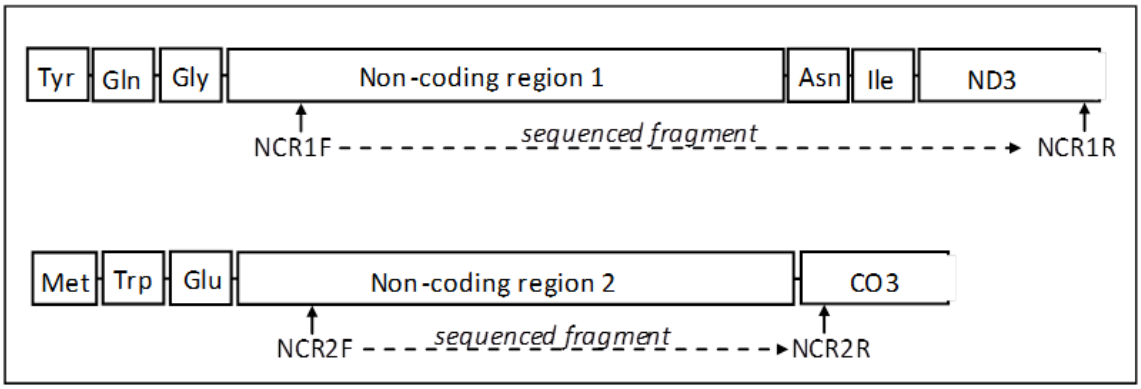


Figure 2.2.1: Position of primers for amplification of non-coding regions 1 and 2 in relation to the full non-coding regions and their flanking genes.

short NCRs (>100bp) present in the mitochondrial genomes of many cephalopod species, that have been shown to be near-identical within individuals (Yokobobri *et al.* 2004). This study aimed to assess sequence variation among duplicated NCRs across three closely related cuttlefish species – *Sepia officinalis*, *Sepia hierredda* and *Sepia vermiculata*. As each species contains two NCRs a first level of analysis was to compare NCR variability between both NCRs within individuals. *S. officinalis* has been identified as an exceptional phylogeographic model species (Perez-Losada *et al.* 2002; 2007), and alongside *S. hierredda* and *S. vermicualta*, the selected taxa represent an elegant continuum of phylogenetic diversity. Accordingly, a primary objective was to investigate the evolutionary modality of each NCR with respect to each other, and the wider mtDNA genome. A second objective was to evaluate the comparative utility of the NCR against traditional mtDNA regions for population genetic, phylogeographic and phylogenetic investigations of the Sepidaee.

2.2.2 Methods

2.2.2.1 Sample acquisition, DNA extraction and PCR amplification

Tissue samples (tentacle clips) of *Sepia officinalis*, *S. hierredda* and *S. vermiculata* were obtained from commercial/artisanal landings as well as through targeted sampling and stored in 95% ethanol for downstream procedures. To increase the likelihood of characterising a large proportion of each species' spatial distribution emphasis was placed on incorporating samples with a comprehensive geographic coverage of each species' distribution.

Table 2.2.1: Location and number of samples employed in comparisons of a) NCR1 & NCR2 within individuals and b) NCR2 with COI and cytb across individuals

		a)			b)		
	Country	code	**NCR1**	**NCR2**	**NCR2**	**COI**	**cytb**
S. officinalis	Turkey, Fethiye	TFE	3	3	11	11	11
	Spain, Alicante	ALI	3	3	16	16	16
	Portugal, Lisbon	PL	3	3	15	15	15
	Bay of Biscay	BB			11	11	11
	English Channel	EC			11	11	11
	Celtic sea	CS	1	1	8	8	8
	Morocco, Essaouira	MES	1	1	14	14	14
	Morocco, Tifnit	MT	3	3	12	12	12
S. hierredda	Senegal, Joal-Faddouth	SEJ	3	3	16	16	16
	Guinea-Bissau	GB	1	1	13	13	13
	Ghana, Tema	GT	3	3	16	16	16
	Angola, Luanda	AL	1	1	2	2	2
	Angola, Kwanza	AK			1	1	1
S. vermiculata	Angola, Luanda	AL	2	2	6	6	6
	Angola, Kwanza	AK			8	8	8
	Angola, Namibe	AN	2	2	5	5	5
	Angola, Flamingo lodge	AF	2	2	4	4	4
	South Africa, Jeffereys bay	SJB	1	1	6	6	6
	South Africa, Bushmans river	SBR	3	3	16	16	16

Table 2.2.2: Primer sequences and location for the amplification of both Non-coding regions of the *Sepia officinalis* species complex.

primer	Primer sequence	Gene location
NCR1-F	5'- TCTTTCTATCTTGCGCCGTGT -3'	Non-coding region 1 (+133bp)
NCR1-R	5'- AAAAGGGGAGCGGGTAGAGA – 3'	ND3 (+332bp)
NCR2-F	5' -TCTTTCTATCTTGCGCCGTGT – 3'	Non-coding region 2 (+110bp)
NCR2-R	5' - TGAGCCGGTTAAAGGTCAAG – 3'	CO3 (+57bp)

Total genomic DNA was extracted from tissue samples using a standard CTAB- chloroform/ isoamylalcohol method (Winnepenninckx *et al.* 1993). Primers were designed to amplify both NCR1 (located between trn-Gly and trna-Asn) and NCR2 (located between trna-Glu and CO3) using a reference sequence of the entire mitochondrial genome of S. officinalis (GENBANK Accession=AB240155.1). For amplification of NCR1 the forward primer was designed to bind to a conserved region at the start of the first NCR, with the reverse primer designed to bind to the ND3 gene. Similarly the forward primer designed to amplify the second NCR was also designed to bind to a conserved region at the start of NCR2, with the reverse primer designed to bind at the start of the CO3 gene (Table 2.2.1; Fig. 2.2.1). PCRs were performed in a total volume of 20 µl, containing 4 µl template DNA, 0.5 µM forward primer (NCR1F for amplification of NCR1 and NCR2F from amplification of NCR2) and 0.5 µM of reverse primer (NCR1R for amplification of NCR1 and NCR2Rfrom amplification of NCR2), 2 mM MgCl2, 0.2 mM dNTP mix (20 µM each dATP, dCTP, dGTP, dTTP), 1x reaction buffer [75 mM Tris-HCl, 20 mM (NH4) 2SO4] and Taq polymerase (BIOTAQ, 5 U/µl). The PCR thermoprofile for NCR1 amplification was: 180 seconds at 95°C, followed by 40 cycles of 30 seconds at 95°C, 45 seconds annealing at 58°C and 60 seconds at 72°C, followed by a final 3 minute extension step at 72°C. For NCR2 amplification PCR conditions were: 180 seconds at 95°C, followed by 35 cycles of 30 seconds at 95°C, 45 seconds annealing at 60°C and 60 seconds at 72°C, followed by a final 3 minute extension step at 72°C. PCR amplicons were cleaned using SURECLEANPLUS (BIOLINE), and sequenced using BigDye technology. For preliminary comparisons of sequences from both NCRs within the same individuals a subset of individuals were sequenced for both NCR1 and 2 (Table 2.2.2): Based on the comparison of patterns between both NCR1 and NCR2 a more extensive screening of was performed for NCR2. For each of these samples comparative COI and cytB sequences were also obtained using methods described in Chapter 2.1.

2.2.2.2 Alignment, quality control and comparisons of NCR1, NCR2, COI and cytb

NCR sequences were inspected to determine sequence quality, and check for the presence of mixed peaks, which may indicate amplification of nuclear copies. Sequences belonging to each NCR were then aligned separately using the CLUSTALW extension, implemented in BI-

OEDIT (Hall 1999). In order to confirm amplification of the correct NCR, sequences of both NCRs were aligned to 2 subsections of the reference genome (GENBANK Accession=AB240155.1) which covered the flanking regions and the full sequences of both NCRs (Fig. 2.2.1). The flanking regions of all sequences successfully aligned with the correct reference sequence, consequently it was considered that both NCRs of the cuttlefish mitochondrial genome had been amplified separately and correctly, and those regions of the sequence that aligned with coding genes/ tRNAs were removed from the alignments.

Across the three species sequence diversity was compared (i) between NCRs and (ii) between NCR2 and the cytB and COI. The distribution of polymorphic sites and conserved regions across each NCR was assessed using entropy (Hx) plots, implemented in BIOEDIT. Haplotype and nucleotide diversities (Nei 1987) were calculated in DnaSP (Librado & Rozas 2009). Finally to assess the power of each locus to resolve inter- and intra- specific phylogenetic patterns, maximum-likelihood phylogenies were constructed in MEGA (Tamura *et al.* 2007), with support calculated after 1000 bootstrap replicates. The optimum substitution model for each marker (NCR2: HKY+G; COI: T92+G+I; cytb: HKY+I) was chosen based on AIC (Akaikine 1974) measures implemented in MODELTEST (Posada & Crandall 1998).

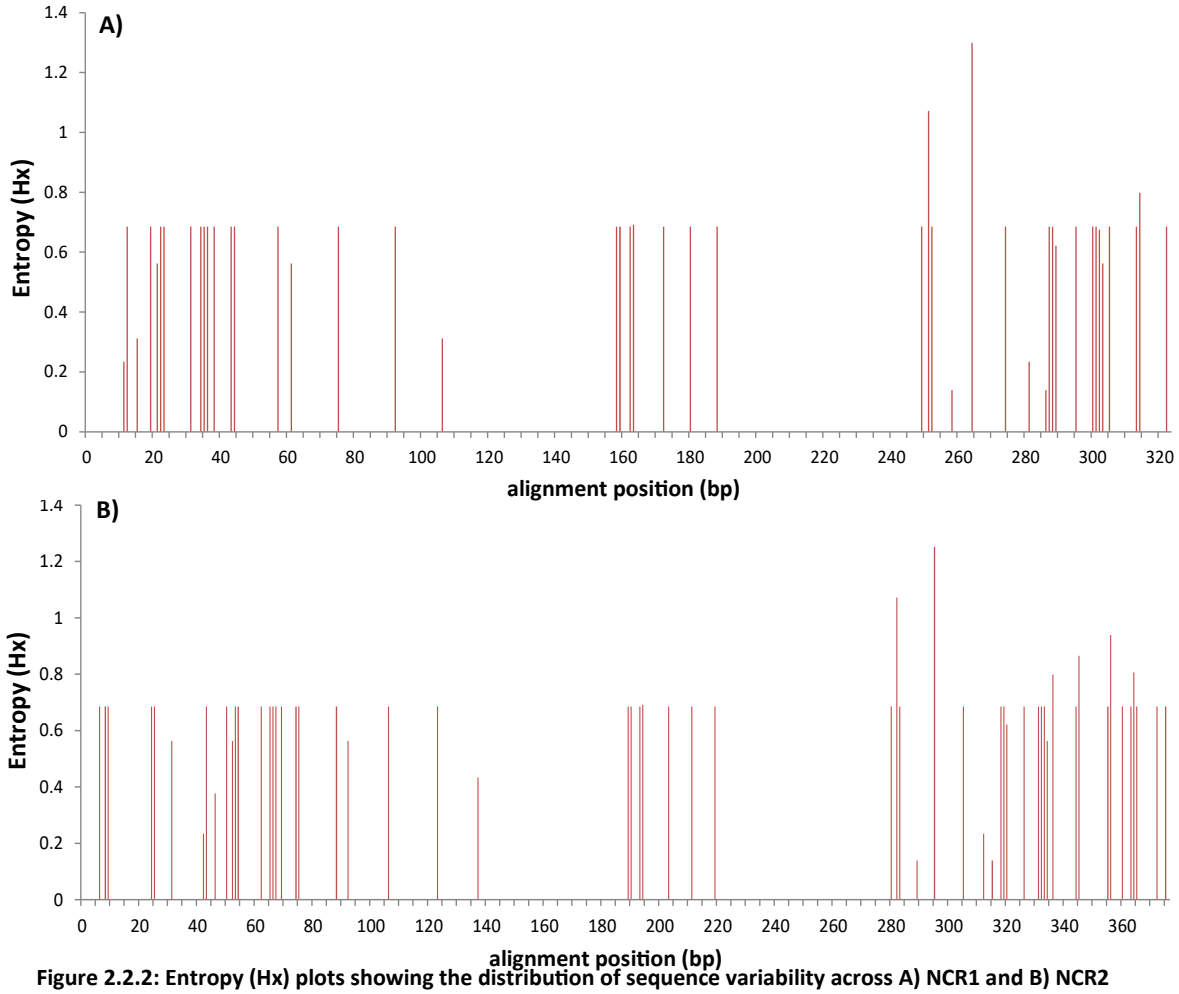


Figure 2.2.2: Entropy (Hx) plots showing the distribution of sequence variability across A) NCR1 and B) NCR2 alignments.

Table 2.2.3: Summary of polymorphism and diversity for NCR1 and NCR2 (long and short sequences) across and within Sepia species. Mono= the number of monomorphic sites, Poly= the number of polymorphic sites, I/G= the number of indels/gaps, PIS=parsimony informative sites, Nhap= number of haplotypes resolved, H= haplotypes diversity (standard deviation in brackets), π= nucleotide diversity (standard deviation in brackets). Number of haplotypes and diversity statistics calculated not considering gaps/indels.

	All species combined (N=32)			Sepia officinalis (N=14)			Sepia hierredda (N=8)			Sepia vermiculata (N=10)		
	NCR1 323bp	NCR2 323bp	NCR2 376bp	NCR1 323bp	NCR2 323bp	NCR2 376bp	NCR1 323bp	NCR2 323bp	NCR2 376bp	NCR1 323bp	NCR2 323bp	NCR2 376bp
Mono	277	278	317	315	314	366	320	320	371	320	322	373
Poly	44	43	55	6	7	7	3	3	4	2	0	1
I/G	2	2	4	2	2	3	0	0	1	1	1	2
PIS	42	41	53	4	5	5	2	1	1	0	0	1
Nhap	14	12	13	7	7	7	4	4	5	3	1	2
H	0.913 (0.032)	0.867 (0.042)	0.909 (0.027)	0.890 (0.050)	0.879 (0.058)	0.879 (0.058)	0.821 (0.101)	0.643 (0.184)	0.786 (0.151)	0.378 (0.181)	0.000 (0.000)	0.467 (0.132)
π	0.0627 (0.003)	0.0617 (0.003)	0.0696 (0.062)	0.0048 (0.001)	0.0055 (0.001)	0.0048 (0.001)	0.0038 (0.001)	0.0032 (0.001)	0.0034 (0.001)	0.0012 (0.001)	0.0000 (0.000)	0.0013 (0.000)

2.2.3 Results

2.2.3.1 Variability between NCRs

In total 323bp of NCR1 and 376bp of NCR2 were amplified across 32 individuals, and no mixed peaks were detected among any sequences. Both NCRs were AT rich. The nucleotide composition of NCR1 was: 40.8% T, 38.2% A, 15.2% C and 5.8% G, and the nucleotide composition of NCR2 was: 39.6% T, 39.7% A, 14.4% C and 6.3% G.

Across all 32 individuals consisting of *S. officinalis (n =14)*, *S. hierredda* (n=8) and *S. vermiculata* (n=10), 27 polymorphisms were observed (Supplementary Table 2.2.1) between an individual's NCR1 and NCR2 (12 transitions, 15 transversions), constituting an average of 99.74% sequence similarity between the two NCRs. The majority of differences resolved between an individual's NCR1 and NCR2 sequences were observed within *S. officinalis*, this included a T-A transversion between NCRs at the tail end of the sequence which was fixed within all *S. officinalis* individuals. An additional 8 nucleotide differences were observed between *S. officinalis* individuals NCRs, resulting in a mean sequence similarity of 99.51%. *S. vermiculata* and *S. hierredda* individuals exhibited a much lower number of differences between NCRs, with 3 nucleotides differing between NCRs in *S. hierredda* and 2 in *S. vermiculata*, resulting in an average sequence similarity of 99.88% and 99.94% respectively.

2.2.3.2 Variability within each NCR

Within each NCR, substitutions were most common at the extremities and central unit of the sequences, with 2 extended conserved regions 51bp (between 107bp and157bp in NCR1 and 138bp and 188bp in NCR2) and 60bp (between 189 and 157bp in NCR1 and 220 and 229bp in

NCR2) in length found at the same location in both alignments (Fig. 2.2.2). In total 44 polymorphic sites were identified across the NCR1 alignment of *S. officinalis*, *S. hierredda* and *S. vermiculata* (Table 2.2.2), with a greater number (55) identified within the full length NCR2 alignment. Although when NCR2 sequences were aligned with and trimmed to the same length as NCR1 only 43 polymorphic sites were observed. Within sequences belonging to *Sepia officinalis* both the short and long fragment of NCR2 contained the most polymorphic sites (n=7), although both NCRs resolved 7 haplotypes. Within *S. hierredda* the long fragment of NCR2 contained the most polymorphic sites (n=4) and also resolved the highest number of haplotypes (n=5). Conversely within *S. vermiculata* NCR1 had the most polymorphic sites (n=2) and reported more haplotypes than NCR2 (Table 2.2.2).

When all species were combined haplotype diversity was slightly higher in NCR1 (H=0.913) than NCR2 (H=0.909), whereas nucleotide diversity was marginally greater in NCR2 (π = 0.0696) than NCR1 (π =0.0627); however standard deviations for estimates of both haplotype and nucleotide diversity overlap across NCRs. When data was partitioned according to species, haplotype diversity was marginally greater in NCR1 (*S. officinalis* H=0.890, *S. hierredda* H=0.821) than NCR2 (*S. officinalis* H=0.879, *S. hierredda* H=0.786) in the *S. officinalis* and *S. hierredda* datasets, whereas in *S. vermiculata* haplotype diversity was greatest for the long fragment of NCR2 (H=0.467).

In order to gain some preliminary insight into the ability of each NCR to detect patterns of inter- and intra- specific divergence, genetic distances between and within species were calculated for both NCRs (Table 2.2.3). Both Tamura-3-parameter and p- distances between species were generally greater for the NCR2 data-set than the NCR1. Within species Tamura-3 parameter genetic distances were identical across NCRs in *S. officinalis* (NCR1/NCR2 = 0.005)

Table 2.2.4: Tamura-3-parrameter genetic distance between (below the diagonal) and within (on the diagonal) *Sepia officinalis, Sepia hierredda* and *Sepia vermiculata*, p distance between species given above diagonal. Standard deviation after 1000 permutations in brackets. A) NCR1 B) NCR2 long

A)	*S. officinalis*	*S. hierredda*	*S. vermiculata*
S. officinalis	0.005 (0.002)	0.177 (0.018)	0.113 (0.018)
S. hierredda	0.295 (0.106)	0.004 (0.003)	0.018 (0.007)
S. vermiculata	0.260 (0.086)	0.022 (0.009)	0.001 (0.001)

B)	*S. officinalis*	*S. hierredda*	*S. vermiculata*
S. officinalis	0.005 (0.002)	0.132 (0.014)	0.125 (0.014)
S. hierredda	0.410 (0.162)	0.003 (0.002)	0.019 (0.006)
S. vermiculata	0.344 (0.132)	0.022 (0.009)	0.001 (0.001)

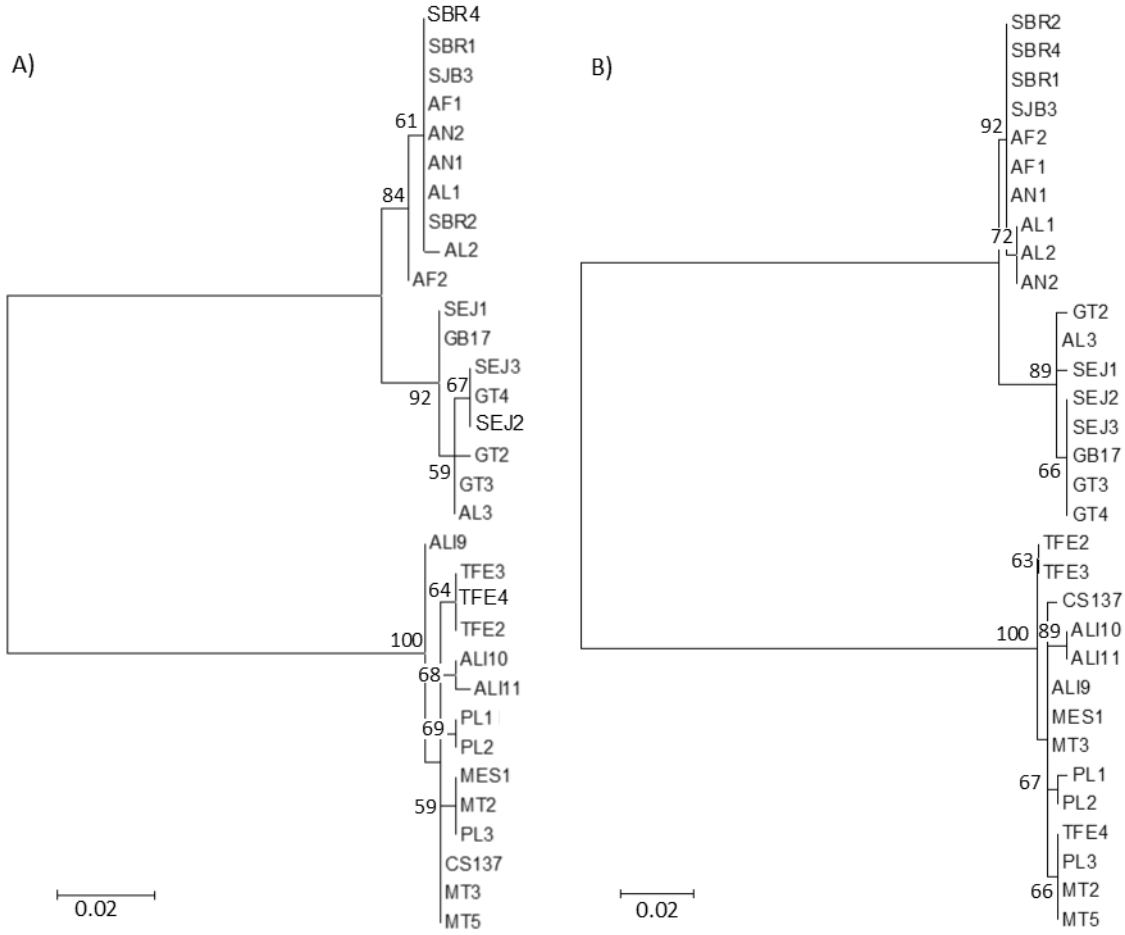


Figure 2.2.3: Unrooted Maximum likelihood topologies based on A) 323bp of NCR1 an B) 376bp of NCR2. Bootstrap support after 1000 replicates given on branches. Sample names correspond to codes given in Table.2.2.1.

and *S. vermiculata* (NCR1/NCR2=0.001), with a slightly greater distance between *S. hierredda* NCR1 (0.004) sequences than NCR2 (0.003)

ML topologies (Fig. 2.2.3) for both NCRs successfully resolved inter-specific relationships, with all three Sepia species forming reciprocally mono-phyletic lineages with high bootstrap support, replicating major divergences resolved using alternative mtDNA markers (Chapter 2.1). Within species however, some disparity was evident between NCR1 and NCR2 topologies. Within *S. officinalis* both NCR1 and 2 consistently resolved three clades comprising of samples from Turkey, Alicante and Portugal, respectively (however in NCR2 TFE4 fell outside of the Turkish clade). Likewise within *S. hierredda* and *S. vermiculata* NCR1 and NCR2 resolved a similar pattern of shallow phylogenetic partitioning, with low BS support.

In conclusion polymorphisms between NCRs were highly conserved (99.74% sequence similarity) and consequently captured nearly equal proportions of each species diversity/ genetic structuring. Both NCRs demonstrated similar power to capture inter-specific relationships, however within species there were slight differences in patterns of genetic diversity and differentiation between NCRs, which varied amongst species.

Table 2.2.5: Summary of polymorphism and diversity for NCR2, COI and cytb across and within Sepia species. Mono= the number of monomorphic sites, Poly= the number of polymorphic sites, I/G= the number of indels/gaps, PIS=parsimony informative sites, Nhap= number of haplotypes resolved, H= haplotypes diversity (standard deviation in brackets), π= nucleotide diversity (standard deviation in brackets)

	All species combined (N=191)			*Sepia officinalis* (N=98)			*Sepia hierredda* (N=48)			*Sepia vermiculata* (N=45)		
	NCR2	COI	cytb	NCR2	COI	cytb	NCR2	COI	cytb	NCR2	COI	cytb
Mono	307	273	410	363	313	478	362	327	483	369	335	494
Poly	64	62	94	10	22	26	12	8	21	5	0	10
I/G	5	0	0	3	0	0	2	0	0	2	0	0
PIS	58	59	86	7	16	16	6	6	13	3	0	8
Nhap	33	28	63	10	18	26	16	9	23	8	1	14
H	0.903 (0.012)	0.885 (0.012)	0.960 (0.006)	0.722 (0.037)	0.837 (0.026)	0.898 (0.017)	0.832 (0.045)	0.682 (0.067)	0.927 (0.020)	0.759 (0.041)	0.000 (0.000)	0.846 (0.043)
π	0.0692 (0.001)	0.0752 (0.001)	0.0670 (0.001)	0.0034 (0.000)	0.0101 (0.001)	0.0062 (0.001)	0.0040 (0.001)	0.0033 (0.001)	0.0055 (0.001)	0.0033 (0.000)	0.0000 (0.000)	0.0037 (0.000)

2.2.3.3 Comparison of the variability and information content of NCR2 with other commonly used phylogeographic markers (COI and cytb)

Ultimately NCR2 was chosen for further phylogeographic/ population genetic analysis of the *S. officinalis* species complex mainly due to the slightly longer length of the amplified fragment. With the primers designed for the amplification of NCR2 reliably amplifying an extra 51bp of NCR sequence, which covered highly variable regions of the NCR (Fig. 2.2.2). As such it was reasoned that when a more comprehensive sample set was screened, the potential to identify novel polymorphisms, and consequently determine patterns of intra-specific diversification was greater for the longer fragment of NCR obtained through amplification of NCR2.

In total 376bp of NCR2, 335bp of COI and 504bp of cytb were amplified across 191 individuals from 17 locations. Despite the NCR/CR generally being considered as a highly polymorphic marker, levels of variability (Table 2.2.5) were consistently greater in cytb. Across all samples 94 polymorphisms and 63 haplotypes were resolved within cytb dataset compared to 64 polymorphisms and 33 haplotypes and 62 polymorphisms and 28 haplotypes in NCR2 and COI respectively. Within each species cytb consistently exhibited a greater number of polymorphisms, however there was some variability between the status of COI and NCR2 as the second most polymorphic marker within species. NCR2 had a higher number of polymorphic sites than COI in *S. hierredda* and particularly in *S. vermiculata* where no polymorphisms and 1 haplotype were observed between all COI sequences within this species. However within *S. officinalis* COI contained over double the amount of polymorphisms as NCR2 and nearly as many as cytb, despite the longer length of the amplified cytb fragment. These patterns were also reflected in measures of haplotype diversity (Table 2.2.5, Fig 2.2.4A), with cytb, again, exhibiting the highest overall (H=0.960) and species specific (*S. officinalis* H=0.898; *S. hierredda* H=0.927, *S. vermiculata* H= 0.846) haplotype diversities, with the lowest haplotype diversi-

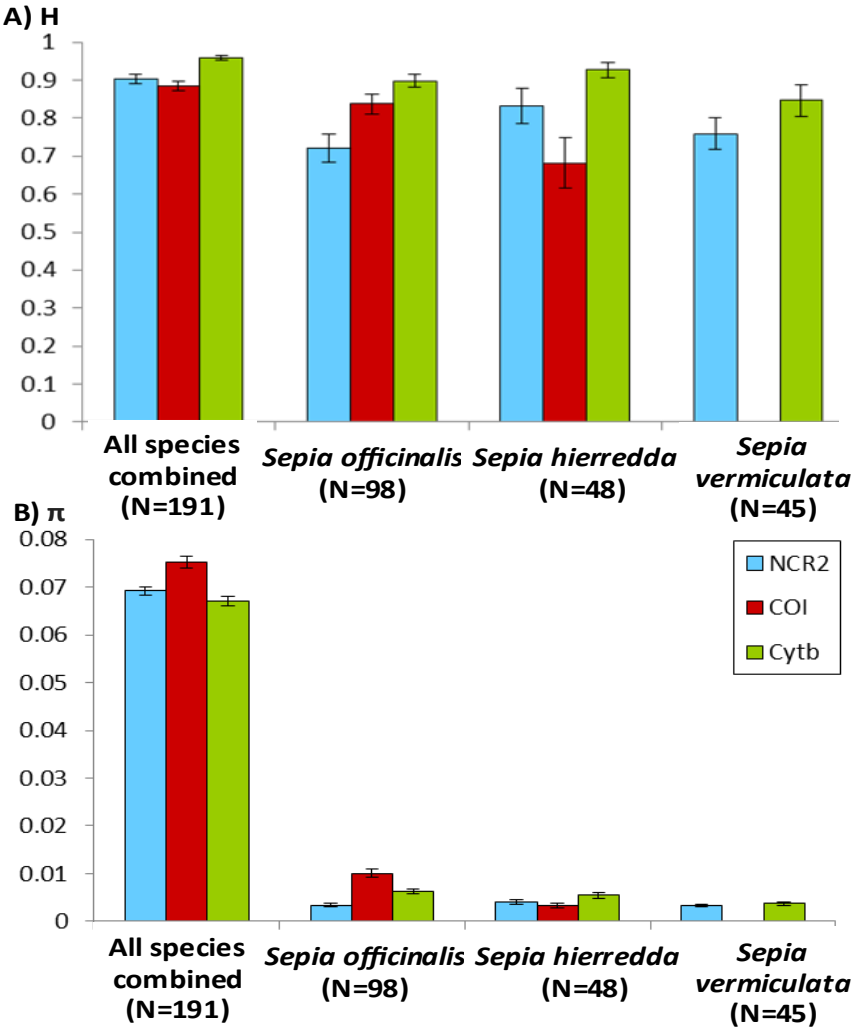


Figure 2.2.4: A) Haplotype (H) and B) nucleotide (π) diversities resolved at NCR2, COI and cytb within *S. officinalis*, *S. hierredda* and *S. vermiculata*, as well as across all species combined.

ty observed at COI in *S. hierredda* (H=0.837) *S. vermiculata* (H=0.000) and the combined datasets (H=0.885) and NCR2 in *S. officinalis* (H=0.722). Conversely for the combined dataset, including all three species, nucleotide diversity (Table 2.2.5, Fig 2.2.4B) was greatest within the COI gene (π =0.0752), and lowest in cytb (π =0.0670). Within species however the cytb locus exhibited the highest nucleotide diversity for both the *S. hierredda* (π =0.0055) and *S. vermiculata* datasets (π =0.0037), with nucleotide diversity within these species lowest in COI (*S. hierredda* π =0.0033; *S. vermiculata* π =0.000). Within *S. officinalis* however COI was the marker with the highest nucleotide diversity (π =0.0101), followed by cytb (π =0.0062) and NCR2 (π =0.0034).

All 3 markers successfully resolved species level divergences between *S. officinalis*, *S. hierredda* and *S. vermiculata*, however ML topologies (Fig. 2.2.5) showed slight differences in the order of these diversifications amongst markers.

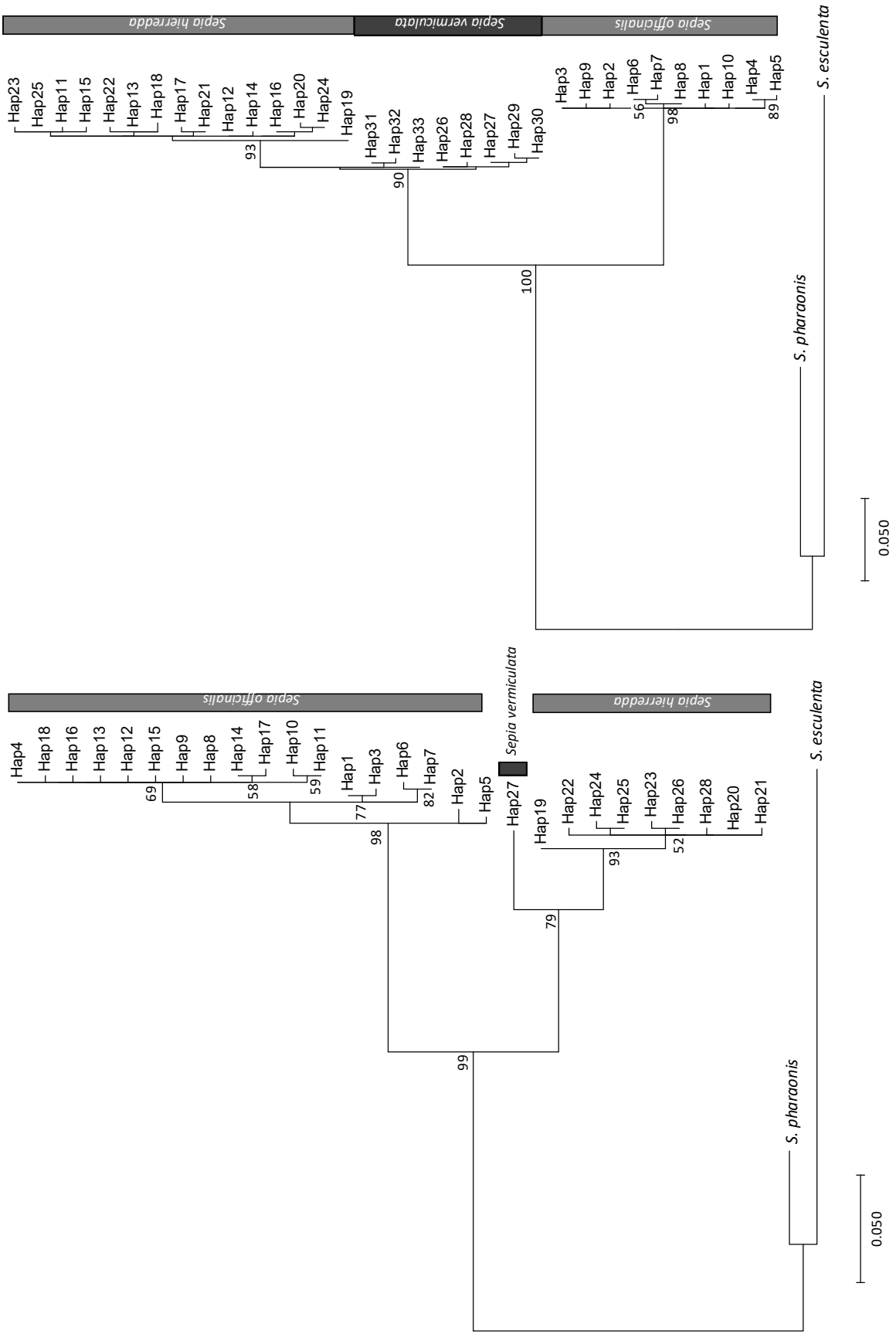


Figure 2.2.5: Maximum likelihood phylogenies of haplotypes resolved from: A) 376bp of NCR2, B) 335bp of COI. Bootstrap support after 1000 replicates given on branches.

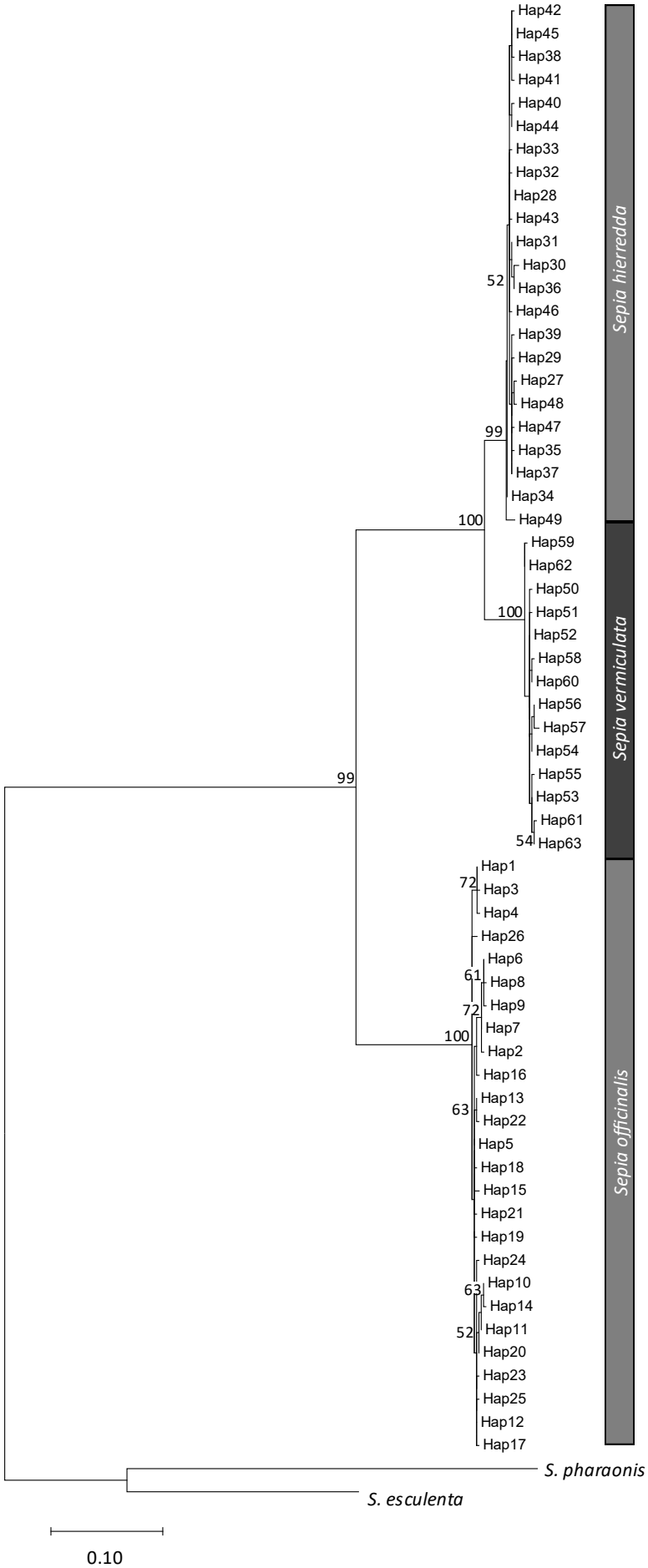


Figure 2.2.6: Maximum likelihood phylogenies of haplotypes resolved from 504bp of cytb. Bootstrap support after 1000 replicates given on branches.

2.2.4 Discussion

The noncoding region has been shown to be rapidly evolving in a range of marine taxa but has been largely unexplored in cephalopods (but see: Akasaki *et al.* 2006; Tomita *et al.* 2002; Yokobori *et al.* 2004; 2007; Winkelmann *et al.* 2013), a group wherein multiple species exhibit duplicated NCRs (Yokobori *et al.* 2004) . This study used the *Sepia officinalis* species complex as a model system for hierarchical sequence analysis with sequence variation first analysed between duplicated NCRs and NCR diversity then compared to that at the more commonly assayed COI and cytb gene regions. There were two salient findings. Firstly there was high sequence conservatism amongst NCRs within species, with both NCRs exhibiting a similar strength to resolve inter- and intra- specific relationships. Second, contrary to expectations that NCR variation would be higher than cytb and COI, cytb revealed the highest levels of divergence with NCR exhibiting similar diversity to COI. As discussed below the results provide information as to the evolution of cephalopod mtDNA genomes as well as the utility of different mtDNA regions for eco-evolutionary analyses.

2.2.4.1 Properties of NCR1 and NCR2

A key finding of this study was that both NCRs exhibited a high amount of sequence conservatism, with sequences of NCR1 and NCR2 sharing on average a 99.74% sequence similarity. One potential explanation for such a pattern is independent and recent duplications within each species as this would explain the homogeneity between sequences within a genome but divergence between species. However, formation of both Sepia NCRs is thought to have been an ancient event occurring early in the evolutionary diversification of the Sepiadae, around 100 million years ago (Akasaki *et al.* 2006). Furthermore, the mtDNA gene order (and thus location of both NCRs) is identical across all three species. Therefore the assumption of independent origin and parallel evolution in each of the three species is much less parsimonious than that of ancestral duplication followed by concerted evolution. Such a dynamic has already been reported for duplicated genes in cephalopods (Tomita *et al.* 2002; Yokobori *et al.* 2004; Winkelmann *et al.* 2013). The mechanisms behind such concerted evolution remain largely speculative, with proposed mechanisms being either some form of inter- or intra- molecular recombination or some selectively beneficial trait associated with replicated NCRs within the mitochondrial genome (Tomita *et al.* 2002; Yokobori *et al.* 2004). Interestingly, the NCRs of *S. officinalis* appear to be less conserved than those of *S. hierredda* and *S. vermiculata*. The higher rate of polymorphisms between *S. officinalis* NCRs may indicate that some ecological/evolutionary factor may be acting to reduce/inhibit the mechanisms that cause conservation of NCRs within this species, although it could also be readily aligned with the

heightened influence of historical isolation and contemporary environmental selection that comes with the greater geographical range of this species.

An additional feature was the high A+T content resolved in both NCRs (NCR1=79%, NCR2=79.3%). Such A+T biases suggest that like in many insect species where similar heightened A+T contents have been resolved (reviewed in Zhang & Hewitt 1997), NCR evolution may be effected by directional mutational selection which has served to conserve this high A+T content (Crozier & Crozier 1993; Jermin *et al.* 1994), which suggests some functional importance of these NCRs.

2.2.4.2 Variability and phylogenetic structuring of the NCR when compared with other mtDNA markers.

A second objective of this study was to assess patterns of variation between NCRs and the more commonly applied cytb and COI regions. Comparing variation among the different regions can provide insight into the roles of forces affecting all regions (demography) and gene specific processes (e.g. hypervariability, secondary structure). A major criticism of the use of mitochondrial NCRs in inter-specific phylogenies is that the typical high mutation and consequent substitution rates of this marker (Chu *et al.* 2003) inhibits the successful resolution of deeper evolutionary splits, and misrepresents phylogenetic diversification (Yang 1998), due to saturation in base substitution and a greater likelihood of inaccuracies and ambiguities in sequence alignment. However, here the NCR, COI and cytb regions all resolved congruent phylogenies that aligned with species biogeography. Specifically, all loci clearly resolved the three species as reciprocally monophyletic and supported the more recent divergence between *S. hierredda* and *S. vermiculata*. Such congruence demonstrates both the influence of demographic forces (e.g. population size, growth or decline and substructure) across the mtDNA genome at this spatio/temporal scale and that these signatures are detectable in the NCR.

The NCR of the mitochondrial genome is typically considered a highly informative and sensitive marker for the shallow (in comparison to aforementioned phylogenetic work) phylogeographic and population genetic analysis of a diverse array of taxa owing in part to its inherently higher mutation rate, when compared with all other regions of the mitochondrial genome (Saccone *et al.* 1987). Within cuttlefish it would appear that the NCR does not exhibit such a higher level of variation than other popular mtDNA markers used in phylogeographic investigations, in particular the cytb gene. Although a high degree of polymorphism is observed in the vast majority of studies that employ NCRs/CRs in marine invertebrates (Chu *et al.* 2003; McMillen-Jackson & Bert 2003; 2004; Grabowski *et al.* 2004; Diniz *et al.* 2005), instances where NCR/CR sequences have failed to resolve expected levels of polymorphism/

diversity have been observed in several investigations of marine mammals (Alter & Palumbi 2009), fish (Meyer 1994; Chubb *et al.* 1998; Tang *et al.* 2006; Page & Hughes 2010) and invertebrate species (Ardnt & Smith 1998; Straughan & Lehman 2000).

This pattern of unexpectedly low NCR variability and higher cytb variability could be caused by both a relaxation of selection on cytb and/or a slowdown in evolutionary rate in the control region. Selection at the interspecies level is evident in the mitochondrial genome of the *S. officinalis* complex (Chapter 2.3) and other species (Rand 2001; Tieleman *et al.* 2009; Scott *et al.* 2010). While low mutation rates at NCRs have been reported for other cephalopods (Winkelmann *et al.* 2013; Aoki *et al.* 2018), it is also possible that mutation rates may appear slow due to hotspot saturation (Galtier *et al.* 2006; Alter & Palumbi 2009) within the NCR. The entropy plots clearly indicate that there is localisation of substitutions in the NCR, and genetic diversity at these hotspots may have reached a maximum level masking the accumulation of additional diversity (mutational saturation).

The slowdown of evolutionary rate at the NCR may also be a product of selective constraints. It has been suggested that genomes with multiple NCRs may be advantageous due to the presence of multiple replication sites providing opportunities for faster replication (Kumazawa *et al.* 1998). Indeed considerably lower levels of variability than expected have been observed in the NCRs of Sepioteithis and Architeuthis squid species (Winkelmann *et al.* 2013; Aoki *et al.* 2018) both of which like the Sepidaee, possess replicated and near identical NCRs in their mitochondrial genome. This loose association between replicated NCRs and lowered sequence variability could indicate that the duplicated NCRs of the cephalopod mitochondrial genome have some unknown functional role. Likely due to the formation of secondary structures between NCRs and as such any mutation within NCRs may be under negative selection, reducing sequence variability. Additionally the two long conserved blocks observed within each NCR (Fig. 2.1.2) may also bear some functional importance and as such may be under heightened evolutionary constraints. Regions of the mtDNA CR have already been shown to contain sequence blocks associated with H-strand termination and origin, alongside both L- and H- strand transcription (Randi & Lucchini 1998; Saccone *et al.* 1991; Sbisa *et al.* 1997); however none of these studies investigated invertebrates. Although to date the regulatory and structural function of cephalopod NCRs remains uncharacterised, the reduced mutation rate in combination with the presence of replicated NCRs across the mitochondrial genome and lengthy conserved sequence blocks within each NCR, suggest that cuttlefish NCRs possess some unknown functional properties, and are likely of greater importance in replication, expression and transcription of the mitochondrial genome than is currently recognised.

The identification of apparent gene specific selection effects in the mtDNA is interesting for a number of reasons. Firstly, it adds to the growing number of studies which suggest non-neutral processes influence mtDNA variability and may thus confound estimates of neutral evolution. More specifically the identification of any form of selection is of interest for cephalopods. Cephalopods are regarded as highly plastic and this plasticity has been suggested to underpin recent climate linked biogeographic changes within the group (Doubleday *et al.* 2016). In comparison to plasticity, genetic adaptations may be too slow relative to current rates of climate change that they may negatively impact species (especially endemic species)/population persistence (King *et al.* 2018).

CHAPTER 2.3

Species level cyto-nuclear discordance indicates introgression and isolation among west African cuttlefish (*Sepia officinalis* species complex)

2.3.1 Introduction

Cephalopods have become an increasingly important fishery resource (Arkhipkin *et al*. 2015; van der Kooij *et al*. 2016) and are viewed as alternatives to many dwindling traditional finfish fisheries (Caddy & Rodhouse, 1998). However, their life histories increase susceptibility to overfishing and some cephalopod fisheries have already declined (Pierce *et al*. 2008). Such declines will have wide-ranging impacts on the marine ecosystems within which cephalopod species occupy important roles as predators and prey (Parry 2006; Xavier *et al*. 2015). Therefore, there is an urgent need to understand connectivity patterns in this group (Allendorff *et al*. 2010). As many ontogenetic tagging methods are not applicable to cephalopods (Arkhipkin 2005) there have been numerous population genetic studies of cephalopod species in different areas. While there are examples of differentiation associated with habitat barriers, the general trend across cephalopods is one of no genetic structuring over wide geographical areas of continuous habitat (Shaw *et al*. 1999; 2010; Reichow & Smith 2001; McKeown *et al*. 2019). In such cases, there is a concern that genetic homogeneity may reflect a lack of resolution of the markers employed to resolve contemporary dispersal restrictions. Furthermore, against such backgrounds of weak/absent genetic structure, many cephalopods exhibit high levels of phenotypic variation indicating high levels of environmental plasticity and/or local adaptation (Van der Vyer *et al*. 2016; McKeown *et al*. 2019). Disentangling such genetic and plastic components is important with respect to predicting species' responses to environmental change and harvesting pressure (King *et al*. 2017; Therkildersen *et al*. 2013), as well as understanding the nature of the recent range shifts observed in many cephalopod species (Arkhipkin *et al*. 2015; van der Kooij *et al*. 2016).

Advances in next generation sequencing technologies, such as Restriction enzyme Associated DNA sequencing (RADSeq) (Baird *et al*. 2008) have dramatically increased the number and type of loci that can be assayed in non-model species. Consequent increases in statistical power have permitted identification of distinct genetic stocks that were not detected in previous studies using traditional population genetic approaches (Mullins *et al*. 2018). In addition, the ability to identify loci potentially under selection (outlier loci) has proven particularly informative in large marine populations (De Wit & Palumbi, 2013) and provided information as to patterns and drivers of local adaptation as well as cost effective tools for routine fish traceability (Nielsen *et al*. 2012).

The objective of this study was to use RADSeq to describe neutral and non-neutral genetic patterns within the *Sepia officinalis* species complex, which comprises *S. officinalis*, *S. hierredda* and *S. vermiculata*. This range-wide genomic study aimed to address specific questions and knowledge gaps emerging from previous research. Specifically, a mtDNA based study (Healey *et al.* 2017; Chapter 2.2) reported three reciprocally monophyletic mtDNA lineages which, alongside biogeographic and morphological information, were ascribed to *S. officinalis*, *S. hierredda* and *S. vermiculata*. However, a surprising finding from that study was the identification of *S. vermiculata* mitotypes, previously considered endemic to South African waters and restricted by the Benguela upwelling system (BUS), in Angolan waters north of the BUS and intermingling with *S. hierredda* mitotypes. This raises two prominent and related questions. Firstly, at the species level, is there reproductive isolation between *S. vermiculata* and *S. hierredda* north of the BUS? Second, at the intraspecific level, does the BUS restrict gene flow between South African and Angola waters for *S. vermiculata*?

In the case of *S. officinalis* previous population genetic studies have reported differentiation at the regional level, i.e. between the Atlantic and Mediterranean and within the Mediterranean basin (Perez-Losada *et al*. 1999; 2007; Turan & Yaglioglu 2010). However, to date none of the microsatellite or mtDNA based studies have detected genetic structure among Atlantic samples. This knowledge gap is pertinent given the known difference in the life history and reproductive characteristics of *S. officinalis* inhabiting this region (Le Goff & Daguzan 1991; Dunn 1999); specifically the two year life cycle of English Channel cuttlefish (Le Goff & Daguzan 1991; Dunn 1991; Royer 2002) compared with the 1 year life span reported in *S. officinalis* across the rest of its distribution. Given this ecological variation across the NE Atlantic, a fundamental aim was to see if the power of RADSeq might detect neutral and/or non-neutral genetic heterogeneity that was beyond the resolution of previous studies, and as such could point to demographic or adaptive divergence and inform management unit delineation.

2.3.2 Methods

2.3.2.1 Sample acquisition and DNA extraction

Tissue samples (arm tip clips fixed in 95% ethanol) were acquired from a combination of artisanal fisheries and targeted sampling across the distribution of *S. officinalis*, *S. hierredda* and *S. vermiculata*. The sampling strategy (see Fig. 2.3.1; Table 2.3.1) was designed to incorporate locations from across the distribution of all three species within the species complex. Total DNA was extracted from arm clips using a phenol/chloroform/isoamyl-alcohol method (Winnepenninckx *et al.* 1993).

2.3.2.2 SNP genotyping

A panel of SNP loci were obtained using tuneable Genotyping-By-Sequencing (tGBS), a modified version of RAD sequencing (conducted by Data2Bio; Iowa State University through the incorporation of genomic DNA digestion with two enzymes (NspI and BfuCI/Sau3AI) reduces the genome and provides an increased number of reads per site (Ott *et al.* 2017). Double digested tGBS libraries were then sequenced on a Life Technologies Ion Proton System, and resulting reads were de-multiplexed, with each read assigned to individual and barcode sequences removed using a custom Perl script (available at https://github.com/orgs/schnablelab; accessed 20 July 2018). Proton adaptor sequences and chimeric reads that harbour internal restriction enzyme sites were removed using Seqclean (sourceforge.net/projects/seqclean). Prior to alignment, the nucleotides from each read were scanned for low quality bases, i.e. PHRED scores <15 (out of 40), using the software Lucy2 (Li & Chou 2004). Each read was assessed in 2 phases: in the first phase sequences were scanned at each end, whilst in the second phase they were scanned using overlapping sliding windows of 10bp.

A *de novo* genome assembly and the resulting contigs were used as a reference genome, and trimmed reads were aligned to this reference genome using GSNAP (Wu & Nacu 2010). Alignments were then scanned for polymorphisms, with a SNP considered homozygote when at least 15 reads supported the genotype and >90% of all reads at that SNP site shared the same nucleotide. A SNP was called heterozygous when each nucleotide variant was reported at least 10 times, with each allele resolved in at least 30% of the total reads. The Minor Allele Frequency (MAF) was set to 0.1.

2.3.2.3 SNP analysis

All SNP loci were tested for linkage disequilibrium in GENEPOP V4.0 (Raymond & Rousset 1995), and deviations from Hardy-Weinberg expectations were assessed across loci and populations using the R package "hwxtest" (Engels 2016). Allele frequencies, observed and expected hetereozygosities, alongside inbreeding co-efficients (F_{IS}) were estimated in Arlequin v3.5 (Excoffier & Lischer 2010). Putative outlier loci (under positive selection) were identified in BAYESCAN v2.1 (Foll & Gaggiotti 2008). Outlier loci were detected using the default Markov Chain Monte Carlo (MCMC) parameters and using the "decisive" criteria under Jeffereys scale of evidence (Jeffereys 1961). Although the concept of balancing selection is well established there are still methodological limitations for its identification (Hansen *et al*. 2010). Therefore we discuss only loci under potentially positive selection. Data was then partitioned into putatively neutral and positive outlier data-sets and each was subjected to all subsequent analyses separately.

Genetic structure was investigated using the Bayesian clustering method implemented in the program STRUCTURE (Pritchard *et al*. 2000) to identify the most probable number of genetic clusters (*K*) (from a range of 1–5) within the data. The analysis was performed both with and without prior sample information (as recommended by Hubisz *et al.* 2009) and with multiple parameter permutations (admixture and correlated allele frequencies, as recommended by Pritchard *et al.* 2000). Each run consisted of a burn-in of 10^6 steps followed by 5 x 10^6 steps with three runs performed for each *K* model tested. Optimal models were assessed using Puechmaille's (2016) estimators for numbers of clusters (MedMeaK, MaxMeaK, MedMedK and MaxMedK) and visual inspection of STRUCTURE bar plots. Genetic structuring was also assessed using the discriminant analysis of principal components (DAPC) implemented in ADEGENET (Jombart *et al*. 2010). Whereas STRUCTURE assigns cluster membership by minimising Hardy-Weinberg and linkage disequilibria within clusters, DAPC employs fewer assumptions and simply maximises differences between groups while minimising differences within groups, with the optimal number of groups identified using the Bayesian Information Criterion (BIC). Genetic differentiation among samples was quantified using global and pairwise F_{ST} values with significance assessed with *P* values following 10 000 permutations in FSTAT (Goudet, 1995). F_{ST} matrices were visualised using principal coordinate analysis in GENALEX. Mantel tests, implemented in GENALEX, were used to test for isolation by distance using the correlation between pairwise F_{ST} and geographical (shortest sea distances) distances between sample sites.

2.3.2.4 MtDNA sequence generation and assessment of deviations from neutrality across the mitochondrial genome.

To complement nuclear data all individuals employed in SNP analyses were amplified for 3 regions of the mtDNA genome: cytochrome b (cytb), cytochrome oxidase subunit 1 (COI) and the second Non-coding Region (NCR2), using the primers and PCR conditions described in Chapter 2.1 & 2.2. Maximum Likelihood and Bayesian inference phylogenies were constructed in PhyML (Guindon *et al*. 2010) and MrBayes (Ronquist & Huelsenbeck 2003) respectively (using parameters described in chapter 2.2).

The potential for non-neutral evolution of the mtDNA genome was assessed through various estimates of ω, calculated by assessing the ratio of non-synonymous to synonymous substitutions (d_N/d_S), as implemented in Datamonkey (Delport *et al.* 2010). However, ω-based approaches have been shown to be biased against even moderately conserved proteins (e.g. Yang & Nielsen 2002; Perez-Losada *et al.* 2006), where if adaptive evolution is influencing these genes it likely only affects a few amino acid sites (Endo *et al.* 1996; Yang *et al.* 2000). Accordingly, to account for heterogeneity in selection across amino acid sites (e.g. site-specific

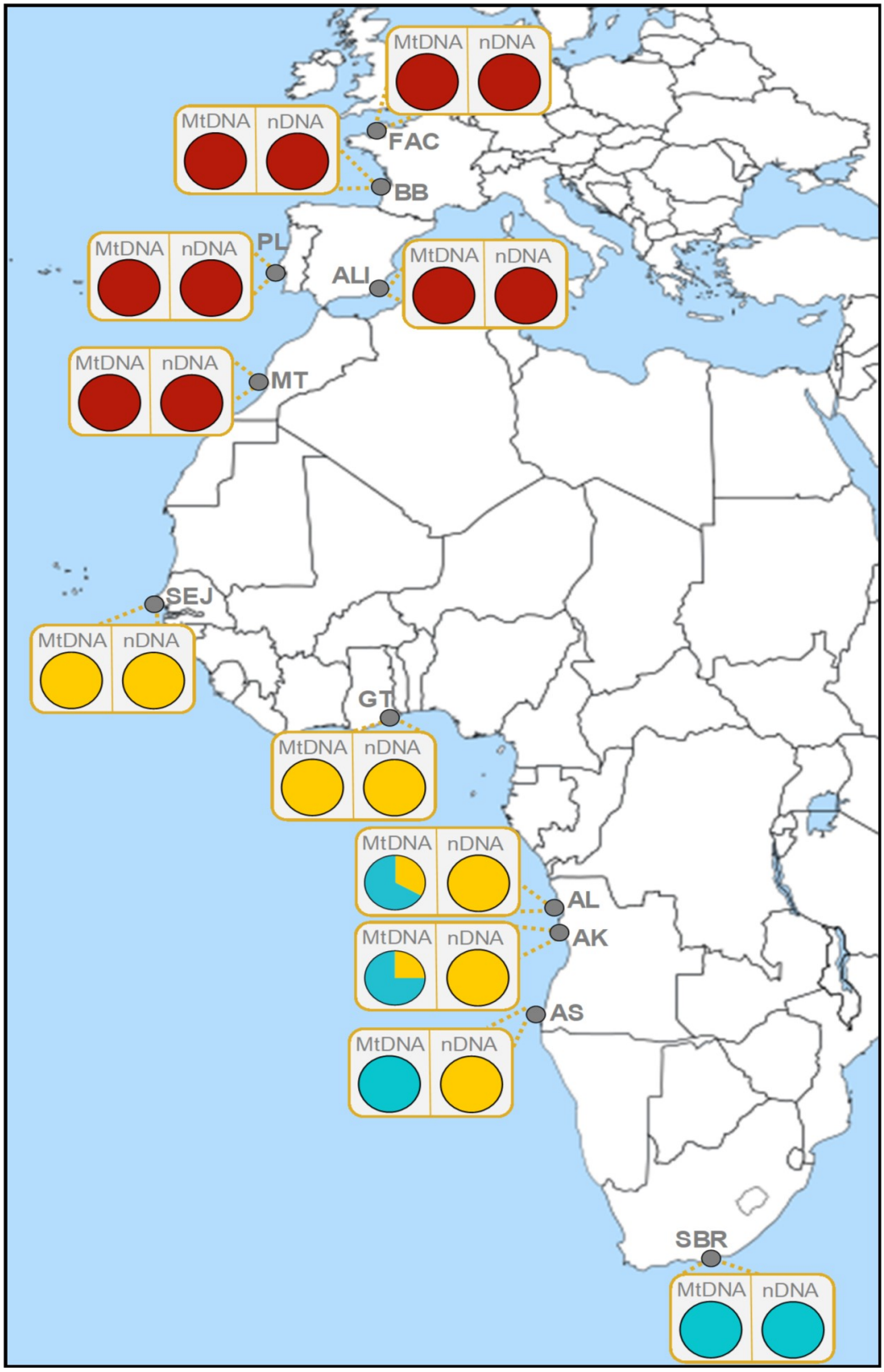


Figure 2.3.1: Sampling locations for *S. officinalis*, *S. hierredda* and *S. vermiculata*, sample codes correspond to Table 2.3.1. Pie charts represent the proportion of individuals within that sample that belong to either the *Sepia officinalis* (red), *Sepia hierredda* (yellow) or *Sepia vermiculata* (blue), mitochondrial (mtDNA) or nuclear (nDNA) types.

differences in adaptive selection), significant functional physicochemical implications of amino acid changes across mtDNA sequences were identified, through implementation of the TreeSAAP (Wooley *et al.* 2003) algorithm. TreeSAAP implements baseml (Yang 1997) to construct a phylogenetic tree, from which the observed distribution of physicochemical changes (defined based on the AA index, Kawashima & Kanehisa 2000) are compared to those expected under neutral theory (e.g. the assumption that all mutations are random and selectively neutral, Kimura 1991). The extent to which amino acid substitutions influence 31 physicochemical properties of that amino acid were estimated based on 8 magnitude categories, however only those that corresponded to the most radical changes (categories 7 & 8) at $p<0.001$ (z score>3.09) were considered, as they are likely to alter protein biochemistry, and as such are likely under strong directional selection (McClellan *et al*. 2005).

2.3.3 Results

2.3.3.1 SNP genetic structuring

A total of 475,818,887 raw reads were sequenced across 85 individuals representative of *S. officinalis*, *S. hierredda* and *S. vermiculata*, with an average of 4,956,446 reads (minimum= 1,145,926, maximum=12,375,552) per individual. After quality control (e.g. filtering of low quality bases and alignment to a *de novo* reference genome) and removal of all SNP loci missing in >10% of individuals, a final SNP dataset consisting of a 6976 SNP loci across 85 individuals was generated. Bayescan analysis of F_{ST} values across all 3 Sepia species (Fig. 2.3.2A) identified 50 putative outlier loci, 39 of which were candidates for positive selection, with the remaining 11 likely under purifying selection. Within *S. officinalis* only, no positive outliers were

Table 2.3.1: Sample information and diversity indices across neutral SNP genotypes. N= the number of individuals genotypes, Npoly= the number of polymorphic SNPs, H_O = Observed heterozygosity, significant ($p<0.05$) deviations from Hardy-Weinberg equilibrium indicated in bold. H_E= expected heterozygosity and F_{IS}. Significance indicated in bold. Standard deviation given in parenthesis.

Sample	Code	N	Npoly	H_O	H_E	F_{IS}
Spain: Alicante	ALI	10	826	0.293 (0.197)	0.303 (0.159)	0.027
Portugal: Lisbon	PL	10	853	0.299 (0.204)	0.310 (0.154)	0.019
Bay of Biscay	BB	10	520	0.305 (0.221)	0.291 (0.161)	-0.023
English Channel	FAC	7	510	0.303 (0.215)	0.308 (0.150)	0.023
Morocco: Tiznit	MT	8	943	0.283 (0.193)	0.290 (0.141)	0.027
Senegal: Joal-Fadiouth	SEJ	10	1170	0.234 (0.176)	0.238 (0.140)	0.020
Ghana: Tema	GT	8	1104	0.243 (0.177)	0.252 (0.136)	0.032
Angola: Luanda	AL	6	954	0.288 (0.197)	0.293 (0.138)	0.010
Angola: Kwanza	AK	4	751	0.349 (0.206)	0.351 (0.124)	-0.007
Angola: Flamingo	AS	2	480	0.520 (0.269)	0.555 (0.098)	-0.020
South Africa: Bushman's river	SBR	10	221	**0.407 (0.308)**	0.329 (0.163)	**-0.176**

detected, with only outliers identified as candidates for purifying selection resolved, even when individuals from the Mediterranean (Alicante), the most divergent genetic group was removed from the dataset (Fig 2.3.2B).

Bayesian clustering of SNP genotypes in STRUCTRE (Fig. 2.3.3B) partitioned the data into 3 clear genetic groups: (1) *S. officinalis* (comprising individuals from Spain-Alicante, Portugal-Lisbon, the Bay of Biscay, the English Channel and Morocco-Tiznit), (2) *S. vermiculata* (south of the BUS), and (3) a group which comprised all individuals north of the BUS and south of the Canary current. This third group therefore contained individuals exhibiting both *S. hierredda* and *S. vermiculata* haplotypes (Fig. 2.3.3). The same three groups were resolved by clustering

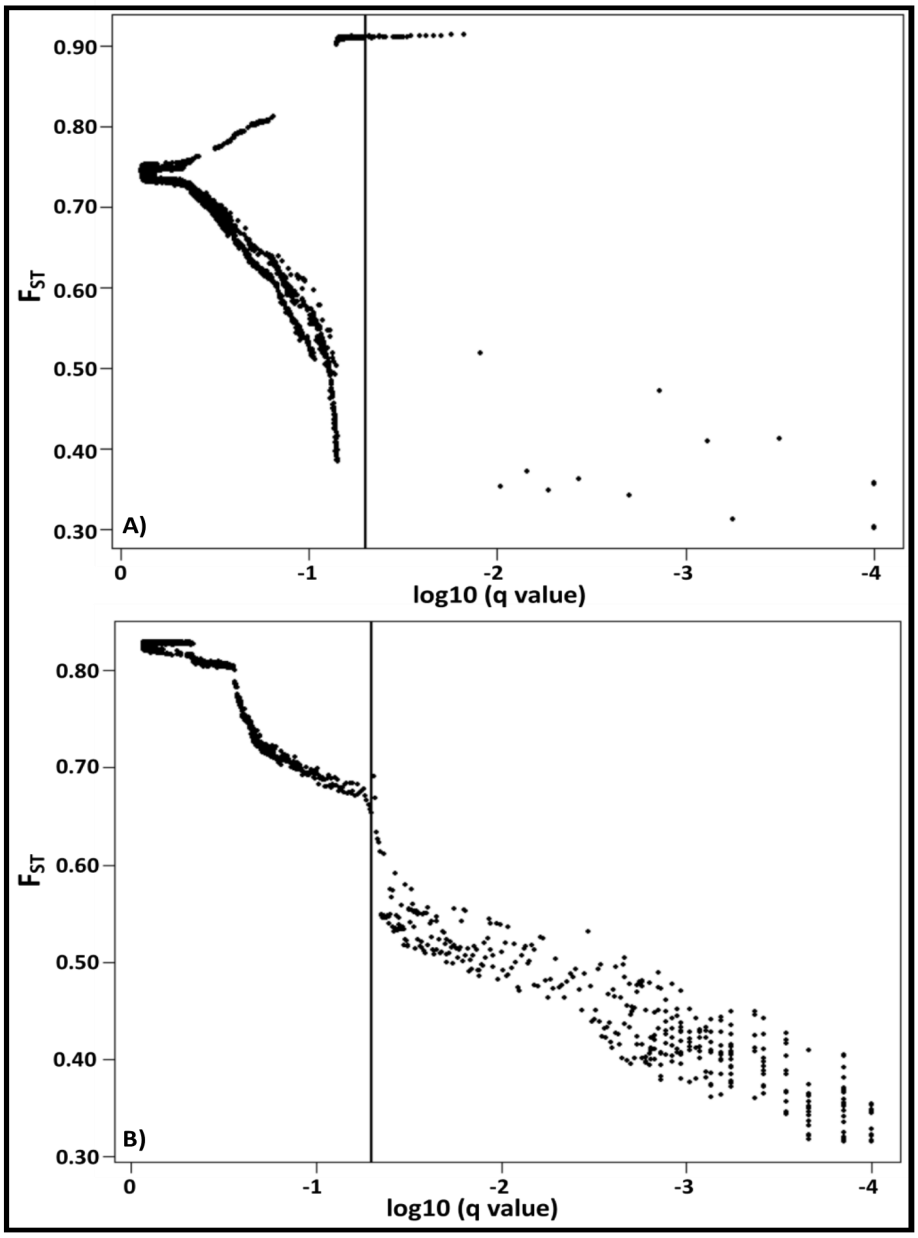


Figure 2.3.2: Outlier loci identified A) across all Sepia species (*Sepia officinalis, Sepia hierredda & Sepia vermiculata*) and B) across Atlantic populations of *Sepia officinalis*. The decision factor to determine selection (log10 q value) is shown on the x axis and locus specific genetic divergence (F_{ST}) on the y axis. The black line indicates a Bayes factor threshold of 95%. All loci with a posterior probability >95% (e.g a log10 value greater than the black line) represent putative outlier loci.

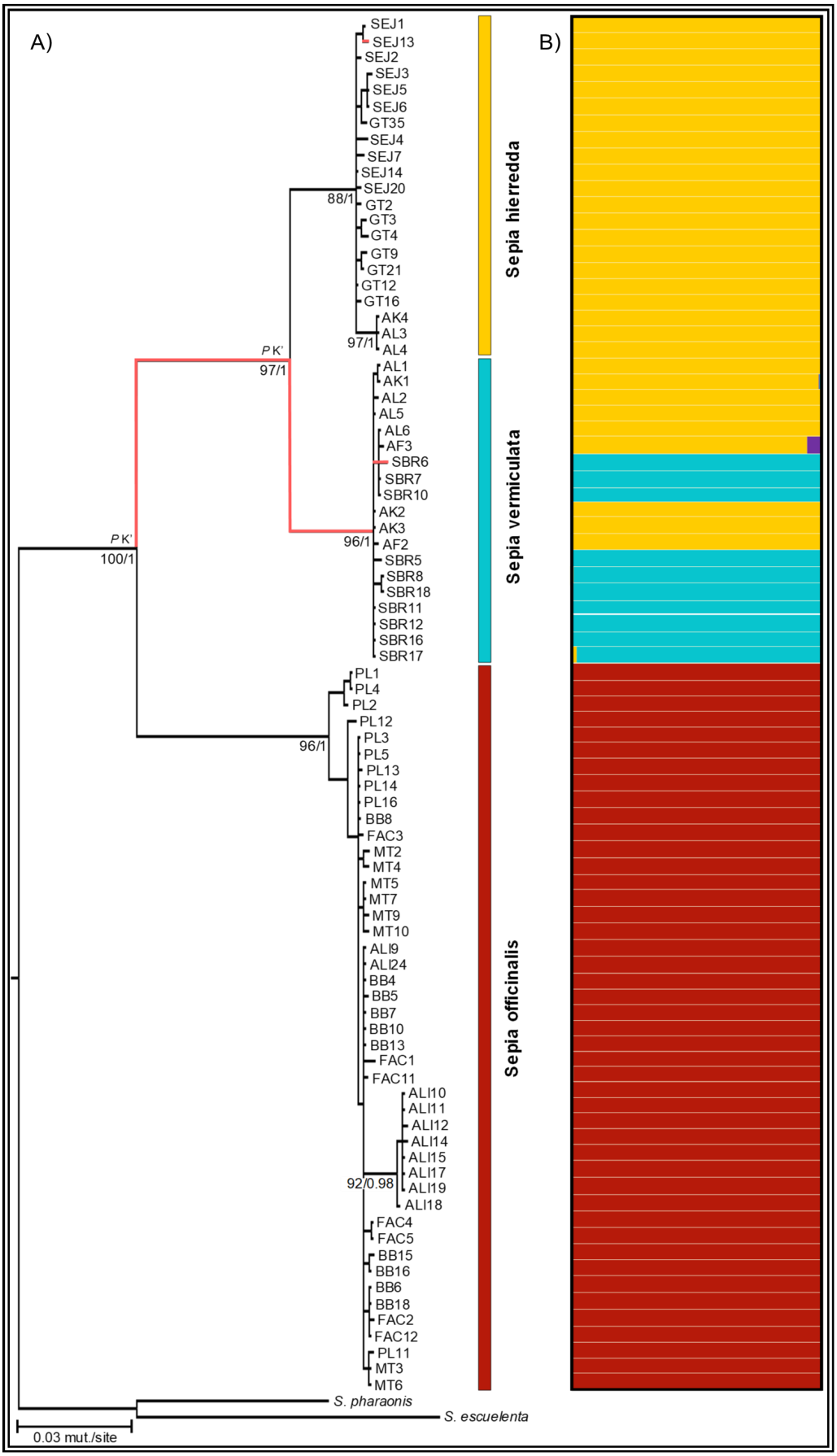


Figure 2.3.3: Phylogenetic/ population genomic partitioning of *Sepia officinalis*, *Sepia hierredda* and *Sepia vermiculata*. A) Maximum likelihood phylogeny based on 839bp of concatenated mtDNA (cytb & COI), ML bootstrap support/Bayesian Inference support given below the branches, coloured branches indicate radical positively selected physiochemical amino acid changes identified in TreeSAAP. B) Bayesian posterior probability of membership to one of 4 (K=4) genetic populations calculated using 6937 neutral SNP genotypes and implemented in STRUCTURE, Each row represents an individual and sample codes correspond to Table 2.3.1.

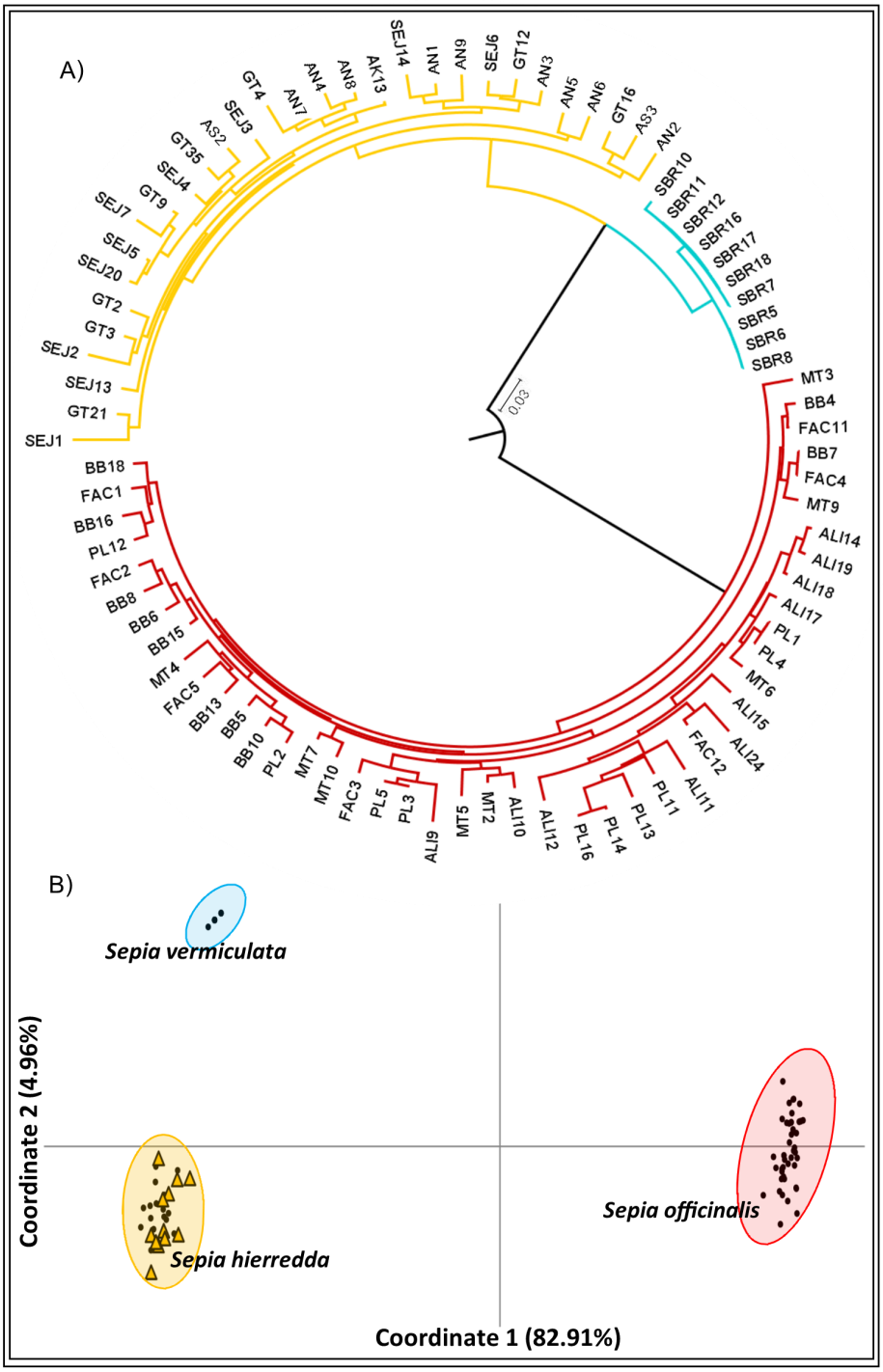


Figure 2.3.4: Distance based clustering based on 39 putative positive outlier SNP loci A) Unrooted Neighbour Joining tree B) Principal Coordinates Analysis, yellow triangle represent individuals from Angola. Red denotes *Sepia officinalis*, yellow *Sepia hierredda* and blue S. vermiculata nuclear types.

analyses of the outlier loci (Fig. 2.3.4) and evident in F_{ST} based analyses of neutral and outlier loci (Table 2.3.2).

2.3.3.2 Deviations from neutral evolution across the mitochondrial genome

TreeSAAP identified several significant (P<0.001, Z score>3.09) and radical (category 7-8) physiochemical amino acid changes among residues in the cytb gene (Table 2.3.3; Fig. 2.3.3A), which indicate strong directional selection. In particular radical (category 8) amino acid replacements which were estimated to influence the equilibrium constant (ionization of COOH) property, were identified on the internal branches separating *S. hierredda* and *S. vermiculata*

Table 2.3.2: Pairwise measures of genetic differentiation (F_{ST}) between samples. Estimates across neutral loci given below the diagonal and estimates based on 39 putative positive outlier loci are given above the diagonal. Statistically significant (p<0.05) values indicated in bold.

	ALI	PL	BB	FAC	MT	SEJ	GT	AL	AK	AF	SBR
ALI		**0.146**	**0.328**	**0.234**	**0.154**	**0.891**	**0.902**	**0.893**	**0.912**	**0.900**	**0.937**
PL	**0.191**		**0.229**	**0.149**	**0.138**	**0.881**	**0.891**	**0.880**	**0.898**	**0.883**	**0.927**
BB	**0.305**	**0.194**		0.007	**0.094**	**0.890**	**0.902**	**0.893**	**0.913**	**0.902**	**0.940**
FAC	**0.295**	**0.178**	0.003		0.060	**0.885**	**0.899**	**0.887**	**0.912**	**0.896**	**0.944**
MT	**0.186**	**0.146**	**0.153**	**0.146**		**0.879**	**0.891**	**0.880**	**0.902**	**0.886**	**0.933**
SEJ	**0.817**	**0.807**	**0.845**	**0.836**	**0.806**		0.007	0.013	**0.096**	0.039	**0.654**
GT	**0.818**	**0.808**	**0.850**	**0.841**	**0.807**	**0.014**		0.007	0.095	0.050	**0.706**
AL	**0.819**	**0.808**	**0.856**	**0.846**	**0.806**	**0.056**	**0.042**		0.004	0.034	**0.683**
AK	**0.821**	**0.809**	**0.865**	**0.856**	**0.807**	**0.053**	**0.039**	0.003		0.085	**0.830**
AF	**0.817**	**0.802**	**0.871**	**0.863**	**0.800**	**0.092**	**0.078**	**0.030**	**0.048**		**0.837**
SBR	**0.887**	**0.878**	**0.918**	**0.921**	**0.884**	**0.495**	**0.511**	**0.525**	**0.571**	**0.635**	

from *S. officinalis* (p<0.001, z score=4.109), and the branch separating *S. hierredda* from *S. vermiculata* (p<0.001, z score=3.201); alongside the terminal branch leading to a haplotype from Senegal (p<0.001, z score=3.181). Radical (category 7) and significant amino acid changes with estimated effect on the average number of surrounding residues of the amino acid were also identified on the terminal branches leading to haplotypes from South Africa, Bushman's' river (p<0.001, z score=5.153) as well as a clade within *S. vermiculata* containing samples from Angola and South Africa (P<0.001, z score=6.055).

2.3.3.3 Population genomic structuring within *Sepia officinalis*

Pairwise genetic differentiation (Table 2.3.2) amongst all north east Atlantic and Mediterranean samples of *S. officinalis* were high and significant at both neutral (F_{ST}=0.146-0.305) and outlier (F_{ST}=0.094-0.328) SNP loci, except for between the Bay of Biscay and the English Channel at neutral (F_{ST}=0.003) and outlier (F_{ST}=0.007) loci and the English Channel and Morocco-Tiznit

Table 2.3.3: Amino acid substitutions across the cytb alignment of *S. officinalis* (SO), *S. hierredda* (SH) and *S. vermiculata* (SV). Red indicates substitutions with significant and radical (category 8) influence on equilibrium constant (ionization of COOH) and green indicates substitutions with radical (category 7) and significant impact on the average number of surrounding residues.

Mitochondrial type			Amino acid substitution									
SO	SH	SV	17	36	37	40	75	80	82	84	85	91
44			L	V	L	V	V	L	I	V	S	V
1			L	V	L	V	V	M	I	V	S	V
	21		L	V	L	L	I	L	I	I	S	A
	1		L	M	L	L	I	L	I	I	S	A
	1		L	V	L	L	I	L	T	I	S	A
		15	L	V	M	L	I	L	V	I	S	A
		4	L	V	M	L	I	L	V	I	C	A
		1	F	V	M	L	I	L	V	I	S	A
		1	F	V	M	L	I	L	V	I	C	A

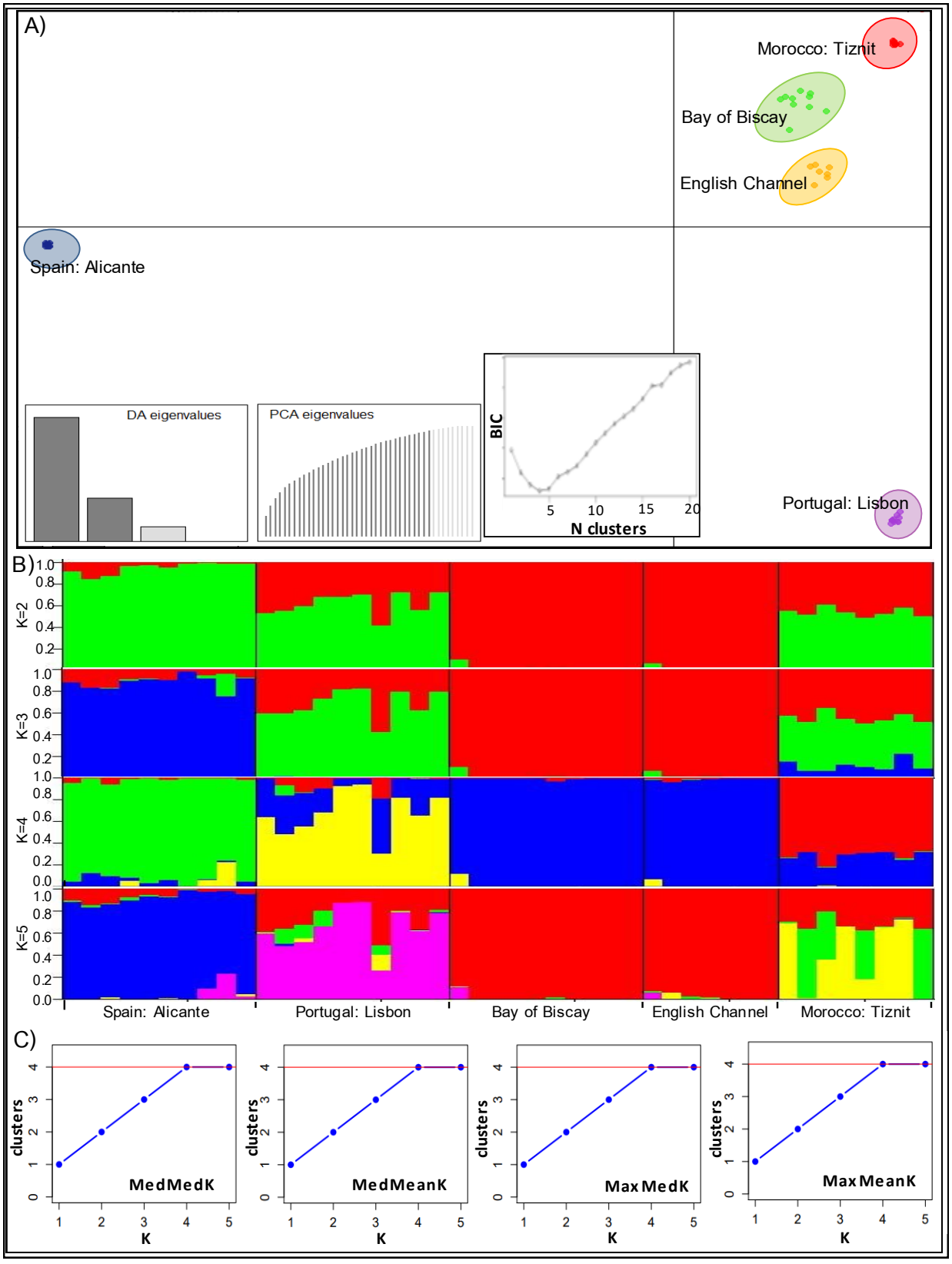


Figure 3.2.4: Clustering of neutral *Sepia officinalis* SNP genotypes calculated using A) Discriminant analysis of principle components and B) Bayesian clustering analysis in STRUCTURE (K=2-5) C) MedMedK, MedMeanK, MaxmedK and MaxMeanK estimates of the most likely value of K in STRUCTURE analyses, with optimal K indicated by the red line.

at just outlier loci (F_{ST}=0.06), where F_{ST} values were low and non-significant. Isolation by distance tests identified a significant ($p<0.001$) correlation between F_{ST} and geographic distance among Atlantic samples. Bayesian clustering of SNP genotypes (Fig. 2.3.5B) supported a model of K=4 with clear delineation of: (1) Mediterranean (Alicante), (2) Portugal-Lisbon. (3) and Morocco-Tiznit and (4) Bay of Biscay and the English Channel. DAPC plots (Fig. 2.3.5A) supported these patterns, alongside indicating the closer clustering of north African (Morocco-Tiznit) individuals with those from the Bay of Biscay and English Channel as well as the distinctiveness of Mediterranean individuals (Spain-Alicante) from all others, which was further corroborated in mtDNA topologies (Fig. 2.3.3A).

2.3.4 Discussion

The use of RADSeq in marine species has provided unprecedented resolution of patterns of historical phylogeography and contemporary dispersal with direct application to conservation and management. After Combosh *et al*.'s (2016) study of Nautilus this represents only the second investigation to employ RADSeq in cephalopods, and the first for any commercially harvested cephalopod. The number of SNPs resolved was several orders of magnitude larger than the number of markers used in previous population genetic studies of cephalopods (e.g. Shaw *et al*. 1999; Perez-Losada *et al*. 2002; Wolfram *et al*. 2006; Sanchez *et al*. 2016), which confirms the potential of this method to generate 1000s of SNPs in this non-model group. This study was also the first to assess nuclear genetic relationships within the *S. officinalis* complex with a specific focus on providing a nuclear context to previous unexpected mtDNA results (Healey *et al*. 2017; Chapter 2.2) which revealed the occurrence of the *S. vermiculata* lineage north of the BUS alongside *S. hierredda*. The SNP data revealed a robust partitioning of three genetically distinct regional groups. One group comprised all the individuals from the *S. officinalis* geographic range (i.e. north of the Canary current) and bearing *S. officinalis* mtDNA haplotypes. Another group comprised all the *S. vermiculata* individuals from south of the BUS. However, individuals harbouring both *S. hierredda* and *S. vermiculata* haplotypes collected north of the BUS all assigned to the third genetic group. This demonstrates both, (i) the genetic cohesion between individuals harbouring *S. vermiculata* and *S. hierredda* mitotypes north of the BUS and (ii) the clear differentiation of *S. vermiculata* south of the BUS.

The genetic differentiation of *S. vermiculata* south of the BUS is not surprising given that the BUS has been shown to be a prominent barrier to gene flow across a range of fish species (Henriques *et al*. 2012, 2014, 2016; Reid *et al*. 2016; Gwilliam *et al*. 2018). However, this is the first demonstration of such a barrier effect for a cephalopod. In a study of *Loligo reynaudii*, Van Der Vyer *et al*., (2016) reported samples north of the BUS to be morphologically distinct from samples south of the BUS. A follow up population genetic study using mtDNA and mi-

crosatellites reported no differentiation across the BUS for *L. reynaudi* suggesting that such phenotypic differences are occurring against a background of high gene flow (McKeown *et al*, unpubl). This inconsistency suggests different cephalopod species may exhibit variable patterns of gene flow and connectivity across the BUS.

A salient feature of the data was that contrary to expectations based on mtDNA co-ancestry (Healey *et al.* 2017; Chapter 2.2), where greater nuclear similarity between *S. vermiculata* descendants either side of the BUS (Angola and South Africa) than to west African *S. hierredda* would be predicted; there was no nuclear differentiation between the *S. vermiculata* and *S. hierredda* clade bearing individuals north of the BUS. This represents a striking case of cyto-nuclear discordance at the species level. Similar discordance between mitochondrial and nuclear genomes has been reported for a number of taxa (e.g. Jacamar species complex, Ferreira *et al.* 2018; brook and arctic charr, Bernatchez *et al*. 1995; Doiron *et al.* 2002; Lamaze *et al.* 2012; chipmunks, Good *et al.* 2008; 2015; and Iberian hares, Seixas *et al.* 2018), with the most common explanation being allopatric divergence followed by secondary contact and introgression. In the present study mtDNA phylogeographic patterns support a scenario of past allopatric divergence between *S. hierredda* and *S. vermiculata*, isolated in regions to the north and south of the BUS, respectively (Healey *et al.* 2017; Chapter 2.2). Following this the contemporary presence of *S. vermiculata* mtDNA haplotypes in Angola is compatible with post-cladogenesis asymmetric dispersal by *S. vermiculata* northward across the BUS. Historical permeability and asymmetric dispersal from South Africa across the BUS into northern waters has been reported for a number of taxa (Reid *et al.* 2016; Gwilliam 2017; Chapter 3.2). A scenario in which the lack of nuclear differentiation between *S. hierredda* and *S. vermiculata* reflects incomplete nuclear lineage sorting in the absence of introgression upon secondary contact can be ruled out as the ratio of between to within clade divergence (approximately 10:1) far exceeds that expected to equate to sufficient evolutionary time needed for the corresponding neutral sorting of nuclear lineages, i.e. the 'three times rule' (Palumbi *et al*. 2001). Accordingly, the data indicate that nuclear introgression in the contact area north of the BUS has eroded nuclear genomic differences that may have existed between descendants of both mtDNA lineages at the time of secondary contact.

Despite the background of nuclear homogeneity there was a clear spatial heterogeneity in the distribution of *S. hierredda* and *S. vermiculata* mtDNA lineages north of the BUS. Specifically, the *S. vermiculata* mitotypes were largely restricted to the cooler and more variable waters at the northern edge of by the BUS, that would have similar environmental characteristics to waters south of the BUS, whereas the S. *hierredda* lineage was more frequent in the more northern tropical waters. This directs inspection as to whether mtDNA may be being shaped by non-

neutral processes. Selection acting directly on mtDNA could generate the observed spatial pattern despite nuclear homogenisation (Bonnet *et al*. 2017). In the case of Sepia a likely scenario is that upon contact and hybridisation leading to nuclear homogenisation of both species north of the BUS, each species mtDNA may have conferred higher fitness within their respective ancestral environment types. In other words, the *S. hierredda* mtDNA type is favoured in the northern more tropical areas whereas the *S. vermiculata* mtDNA type is favoured (and so retained against the swamping nuclear introgression from a much larger resident *S. hierredda* population) in the cooler and variable southern Angolan waters. Although speculative there is growing awareness that selection on mtDNA can drive divergence through metabolic adaptations to different environmental conditions such as temperature, habitat and resource availability (e.g. Rand 2001; Tieleman *et al*. 2009; Scott *et al*. 2010). Furthermore, TREESAAP analysis provided evidence of functional differences among the differing species clades. Although signatures of mtDNA selection were most pronounced in comparisons involving *S. officinalis*, analysis of further gene regions could reveal similarly strong signatures between *S. hierredda* and *S. vermiculata*. Other studies of cephalopod mtDNA have also provided indirect evidence of selection pressures on mtDNA evolution (Akasaki *et al*. 2006; Winkelmann *et al*. 2013).

S. officinalis is an intensively harvested species and as such aligning spatial management units with patterns of population structure is important to ensure sustainability (Reiss *et al*. 2009). Although sample sizes were small, RADSeq revealed a hierarchical pattern of structure within *S. officinalis* which aligned with previous studies and provides new information. The major genetic divergence was between the Mediterranean and Atlantic. This genetic break was previously reported for *S. officinalis* using mtDNA (Perez-Losada *et al*. 2007) and microsatellites (Perez-Losada *et al*. 2002) and has been attributed to vicariance in distinct Atlantic and Mediterranean glacial refugia, with contemporary gene flow being physically restricted by the Almeria-Oran frontal system. Among the Atlantic samples, three spatially coherent genetic groups were resolved: (1) English Channel/ Bay of Biscay, (2) Portugal, and (3) north west Africa. The lack of genetic differentiation between the English Channel and Bay of Biscay aligns with previous microsatellite-based studies (Wolfram *et al*. 2006; Mc Keown pers comm) with individuals from both regions suggested to mix in overwintering areas in the Hurd deep. From a fishery management perspective, the discrepancy between levels of gene flow needed to limit genetic differentiation and dispersal needed to replenish stocks (Hauser & Carvalho 2008) is an important consideration here and thus more individuals will need to be genotyped to ascertain the level of dispersal between both areas.

Across the Atlantic samples there was a significant correlation between genetic and geographical distance which supports the biological significance of the results (Waples, 1998) However,

that does not necessarily mean there is a correlation between contemporary connectivity and geographical distance *per se* (Garnier *et al*. 2004). Correlations between genetic and geographical distances can be artefacts of barrier effects, i.e. localised rather than clinal reductions in gene flow. Hydrographic barrier effects have been suggested to explain genetic breaks between the Iberian coast and Bay of Biscay in a number of studies (Piñeira *et al*. 2008; Woodall *et al*. 2011) and may also restrict gene flow between Iberian waters and the Bay of Biscay in *S. officinalis*. Barriers to gene flow between African and European coasts have also been identified in sessile (Quintero *et al*. 2007) and highly dispersive (Aboim *et al*. 2005) taxa and may similarly be involved in driving the genetic isolation of the north west African population from populations on the European Atlantic coast. However, finer scale sampling will be required to more fully examine the drivers of the observed population structure.

RADSeq offers the opportunity to identify loci potentially under selection which may then provide information as to patterns and processes of local adaptation (Milano *et al*. 2014), ontogenetic dispersal (Mullins *et al*. 2018), as well as offering tools for fish traceability (Nielsen *et al*. 2013). Here multiple marker based neutrality tests identified 39 positive outliers in global tests performed over the three nuclear groups (i.e. *S. officinalis, S. vermiculata* south of BUS and the north of the BUS group) with no positive outliers identified in comparisons restricted to *S. officinalis* only. Such a proportion (0.56%) of outliers are lower than reported in other studies of marine species (0.99% Milano *et al*., (2014); 2.6% Bourret *et al*., (2013), 4.59% Hess *et al*., (2013)). A number of factors have likely contributed to such a low proportion of outliers. Firstly, the high level of neutral divergence between the species groups would undoubtedly bias the number of positive outliers detected compared to a weakly structured system. Confounding factors due to the high level of neutral structure would also contribute to the high number of loci identified as being potentially under balancing selection. Second, although the total number of SNPs identified compared favourably to other studies, the large genome sizes of Sepia species means that the number of loci resolved here may provide insufficient density to detect a representative proportion of SNPs that are actually linked to non-neutrally evolving regions (Lowry *et al*. 2017). Thirdly, in the case of *S. officinalis* the small sample sizes for the different sites may have also limited the power of the multiple marker based neutrality tests employed here. However, in this context it is important to note that the BAYESCAN method has been shown to be the most conservative and to have the lowest Type I error rates of the various F_{ST} based outlier tests currently available (Narum & Hess 2011). In conclusion, the low proportion of outlier SNPs identified at the inter-specific and intraspecific levels of divergence should not be taken as indicative of a lack of non-neutral structuring within the *S. officinalis* species complex when compared to other groups.

Increased genetic divergence across samples at outlier loci might be explained as an effect of directional selection with differentiated groups delineated by such loci corresponding to locally adapted units. Here outlier loci resolved the same three genetic groups as the neutral markers. In light of the clear regional partitioning of these groups and associated divergent environmental regimes that they occupy, adaptive divergence would *a priori* be considered likely. However, demonstration of the adaptive significance of the outlier SNPs would require functional tests which are difficult to perform in non-model marine species. Bierne *et al*., (2011) also posit that outlier loci may reflect endogenous genomic incompatibilities rather than environmental adaptation. Distinguishing between exogenous (adaptation) and endogenous (genomic incompatibilities) is extremely difficult as they are not mutually exclusive and both processes may act concurrently. As such incompatibilities typically evolve during prolonged isolation it is interesting that no nuclear outliers were detected between the *S. vermiculata* and *S. hierredda* clade descendants north of the BUS that could suggest an absence of such nuclear incompatibilities. However, in this context it is important to highlight that incompatibilities could occur at the mito-nuclear level and permit nuclear gene flow while constraining mtDNA gene flow (Beekman *et al*. 2014) and as such explain the decoupled mtDNA and nuclear structuring north of the BUS.

This study has important implications for management and conservation of cephalopod biodiversity. At the intraspecific level the data provides information useful to the spatial management and identification of *S. officinalis* stocks. At a species level the finding of unrestricted gene flow between *S. hierredda* and *S. vermiculata* upon secondary contact is striking in light of predicted and ongoing cephalopod range shifts (Zeidberg & Robison 2007; Ruiz-Cooley *et al.* 2013; Golikov *et al.* 2013; 2014; Ramos *et al*. 2014; Xavier *et al*. 2016). Such range shifts may result in previously allopatric lineages / species coming into contact. Hybridisation amongst species can have serious biodiversity implications (Allendorf *et al*. 2001; Seehausen *et al.* 2008), leading to loss of local adaptation (Laikre *et al.* 2010), extinction of one or both parental species (Allendorf *et al*. 2001) and reverse speciation / species collapse, leading to hybrid swarms, as documented extensively in whitefish (*Coregonus spp*.-Vonlanthen *et al*. 2012; Bhat *et al*. 2014), three spined sticklebacks (*Gasterosteus aculeatus* species complex- Velema *et al.* 2012; Taylor *et al.* 2006) and cichlids (Seehausen *et al.* 1997). Conversely the exchange of genetic information amongst species can also drive the formation of new species, facilitating adaptive radiations that exploit introgressed genetic material (Sternkopf *et al*. 2010; Masello *et al.* 2011) and may allow for novel adaptive evolutionary trajectories under environmental changes (Rieseberg *et al.* 2007; Hedrick 2013; Tigano & Friesen 2016). Therefore, cephalopod responses to climate change may be much more complex than the simple single species move, adapt, acclimatise or die scenarios predicted for other groups.

CHAPTER 3

Evolutionary history of Mediterranean, eastern Atlantic and south-west Indian Ocean grey mullets

CHAPTER 3.1

Preliminary phylogenetic analysis of west African, southern African and Mediterranean Mugilidae reveals discordance between species designations within publicly available barcode databases and those based upon mitochondrial lineage.

3.1.1 Introduction

Grey mullets (Mugilidae) are a widespread and abundant family of fish in West African, Mediterranean and South African waters, wherein they support artisanal and commercial fisheries as well as aquaculture (Harrison & Senou 1999; Bernadon & Vall 2004; Guillemot *et al.* 2009; Jollit *et al.* 2010; Crosetti 2016). In West Africa Mugilidae are recognised as a major economic resource (Brulhet 1975; Payne 1976; Bernardon & Vall 2004), representing 18% of the total landings of the Mauritanian fishery (Bernardon & Vall 2004) and supporting many coastal communities. However many mullet fisheries across West Africa are unregulated (Koranteng *et al.* 2000), and as such their importance as an artisanal resource may be underestimated. In addition species-specific fisheries data are available for few Mugilidae species (Chabanet & Durville 2005; Veiga *et al.* 2010; Franco *et al.* 2012; Mellin *et al.* 2016), which could in part be attributed to the conservative morphology of the Mugilidae hindering identification to species or even genus level by traditional morphological and morphometric methodologies (Schultz 1946; Thomson 1997; Harrison & Senou 1999; Caldara *et al.* 1996).

The lack of reliable species- and/or genus-specific distinguishing morphological features has not only restricted the appropriate management of grey mullets but also limited understanding of the diversity of Mugilidae and their systematic relationships. This morphological conservatism has led to a wide range of anatomical (e.g. Schultz 1946; Senou 1988; Harrison & Howes 1991; Thomson 1997; Ghasemzadeh 1998) and karyotypic (e.g. Cataudella *et al.*, 1974; Delgado *et al.*, 1992; Crosetti *et al.*, 1993; Rossi *et al.*, 1997, 2000) features being used to try to delineate species, and to elucidate the evolutionary history of grey mullets. Perhaps as a result of this methodological variation amongst studies, morphologically and anatomically derived systematic relationships amongst mullet species and genera, and indeed the classification of these species and genera, are contradictory amongst studies (reviewed in Durand *et al.* 2012a). As such a robust taxonomy of Mugilidae has yet to be resolved, although general consensus recognises 14 to 20 genera within the family (Thomson 1997; Ghasemzadeh 1998; Nelson 2006).

A key technological advancement that has contributed greatly to our understanding of mullet taxonomy and evolution is the application of molecular genetic methodologies. Molecular

DNA-based methods can provide insights into the phylogeny, systematics and taxonomy of a species, genus or family; independent from morphology (Pons *et al.* 2006; Craig & Hastings 2007; Rogers & González 2010; Ladner & Palumbi 2012). As such these approaches can be invaluable to investigations of species groups (like the Mugilidae) for which classical phenotype-based methods have failed to reach a consensus in regards to species/genus delineation and evolutionary relationships. However, as the vast majority of molecular taxonomic and phylogenetic investigations of grey mullets have focussed on relatively confined geographic regions relative to their global distribution (see Table 3.1.1) It was not until relatively recently that a molecular insight was provided on global mullet evolution (Durand *et al.* 2012a,b).

Durand *et al.*'s (2012a) global mullet phylogeny employed broad geographic and taxonomic sampling, utilising 3 mitochondrial DNA (mtDNA) regions (COI, cytb and 16s) to investigate the systematics of 19 genera and 55 species. Along with confirming the monophyly of the Mugilidae family, Durand *et al.* (2012a) identified 7 major clades within the family. Moreover, they found many inconsistencies between the major morphological phylogenies (Schultz 1946, Senou 1988, Harrison & Howes 1991, Thomson 1997, Ghasemzadeh 1998) and those based on mtDNA; in particular Durand *et al.* (2012a) suggested that all genera with representatives from 2 or more species (Liza, Moolgarda, Myxus, Oedalechilus, Rhinomugil and Valamugil) were para- or poly-phyletic. Consequently this led Durand *et al.* (2012b) to propose a revision of Mugilidae taxonomy recognising 25 genera, 15 of which were previously described (Agonostomus, Aldrichetta, Cestraeus, Chaenomugil, Chelon, Crenimugil, Ellochelon, Joturus, Mugil, Myxus, Neomyxus, Oedalechilus, Rhinomugil, Sicamugil and Trachystoma), with the remaining resurrected (Gracilimugil, Minimugil, Osteomugil, Planiliza, Plicomugil and Squalomugil) or new (Neochelon, Parachelon and Pseudomyxus). Building upon these investigations, Xia *et al.* (2016) complemented previous mtDNA phylogenies with nuclear markers and morphometric characters, corroborating the proposed genus level designations of Durand *et al.* (2012b), alongside producing a key to morphologically distinguish amongst

Table 3.1.1: Previous phylogenetic/ phylogeographic investigations of Mugilidae and the geographic region which they focussed upon.

Geographic focus of pre-	references
Mediterranean	Autem and Bonhomme (1980), Blel *et al.* (2008), Caldara *et al.* (1996), Erguden *et al.* (2010), Gornung *et al.* (2007), Imsiridou *et al.* (2007), Murgia *et al.* (2002), Papasotiropoulos *et al.* (2001, 2002, 2007), Rossi *et al.* (1998a, 2004), Semina *et al.* (2007), Turan *et al.* (2005)
East Asia	Lee *et al.* (1995), Liu *et al.* (2010); Shen et al. 2011
Americas	Fraga *et al.* (2007); McMahan *et al.* (2013)
India	Menezes *et al.* (1992)
Global	Durand *et al.* (2012a,b), Durand & Borsa (2015), Xia *et al.* (2016)

these genera.

The systematic and phylogenetic uncertainty of the Mugilidae has also been found to extend to the within-species level. Recent phylogenetic investigations that included broad geographic sampling and included samples from across species distributions (Durand *et al.* 2012a,b, McMahan *et al.* 2013, Durand & Borsa 2015) found congruent patterns of independent, deeply diverged and, on occasion, paraphyletic lineages within many mullet species. By 2015 a total of 63 distinct and monophyletic lineages had been identified within the Mugilidae (Durand & Borsa 2015), of which 15 and 6 were found within *Mugil cephalus* and *Mugil curema* respectively. The extreme diversity within these latter two species may in part be attributed to their presumed global or inter-oceanic distributions, but may also be due to the greater phylogeographic focus on these species in comparison with other grey mullets. Whether all distinct DNA lineages within grey mullets represent cryptic species has yet to be confirmed, with new species/sub-species designations only partially recognised within *M. cephalus* (Whitfield *et al.* 2012). Nonetheless to date the frequency of discovery of distinct lineages in relation to the number of currently recognised species highlights that a large proportion of the species richness and diversity within the Mugilidae may remain uncharacterised.

Misidentification and/or inaccurate characterisation of diversity within species has important implications for the conservation and sustainable management of a species, genus or family. Accurate definition of species distributions are central components for management, and a lack of clarity with regard to species richness and distribution can influence estimates of abundance, genetic diversity and genetic differentiation, increasing the chances of mismanagement alongside heightening the risk of local extinctions. In the present study the initial aim was to conduct a comprehensive phylogeographic / population genetic investigation of *M. cephalus* (Linnaeus 1758) and *Chelon dumerili* (Steindachner 1870) populations across their distributions on Mediterranean, west African and southern African coasts, in order to contribute information towards future sustainable management of these species. Although samples were initially targeted and identified morphologically (with the help of local scientists) as either *M. cephalus* or *C. dumerili*, when these samples were genotyped at the mtDNA COI locus a high proportion of sequences failed to cluster phylogenetically with either *M. cephalus* or *C. dumerili* database standards, but rather fell within other distinct mtDNA lineages corresponding to other recognised Mugilidae species described by Durand *et al.* (2012a,b). The high frequency of misidentification within the samples collected here was attributed to the taxonomic ambiguity of the Mugilidae, as well as the lack of obvious morphologically distinguishing features between Mugilidae species. This observation was further supported when compiling data from publicly accessible barcoding databases for the subse-

quent phylogeographic study (chapter 3.2), where numerous instances of barcode misassignment (where the DNA sequence attached to a species designation failed to cluster phylogenetically with the corresponding mitochondrial lineage according to Durand *et al.* 2012a,b) were observed. Consequently the initial aims of the grey mullet study were changed to first provide an overview of publicly available COI barcoding data for all grey mullet species known from South African, west African and Mediterranean coastal waters, supplemented with the novel sequences generated here. Accordingly the main aims of this part of the study were: 1) the identification of anomalous sequence records that do not conform with current species or genus designations, in order to generate an unambiguous sequence dataset as a resource moving forward; and 2) to further resolve evolutionary relationships among these mullet species by including a more comprehensive sequence set than previous phylogenetic investigations of Mediterranean, west African and South African Mugilidae and estimating the relative timing of major diversification events.

3.1.2 Methods

3.1.2.1 Sample acquisition

Samples of grey mullet were collected from a combination of commercial and artisanal fish markets, as well as targeted fishing of juveniles in estuarine habitats, using procedures covered under local ethical and fishing regulations and alongside local staff with suitable permissions. In order to cover the distributions of the target species for this thesis, sampling was conducted from one Mediterranean location (Fethiye, Turkey), two north east Atlantic locations (Olhao, Portugal and Essaouria, Morocco), three tropical west African localities (Tema, Ghana; Kwanza River, Angola and Flamingo Lodge (Namibe), Angola) and six locations across South Africa (SA) to target the cool temperate west coast (Bergirivier lagoon), warm temperate Eastern Cape coast (Kneisner lagoon, St. Francis, Kariega river mouth, Kowie river mouth) and the tropical/subtropical Indian Ocean coast (Mpambanyoni river, Kwazulu-Natal). Although attempts were made to morphologically identify samples to genus/species level, owing to the morphological conservatism of this family it was decided that for the purpose of this investigation and pending barcoding, all samples would be loosely categorised as belonging within the Mugilidae family. Fin clips were taken and fixed in 95% ethanol for downstream procedures.

3.1.2.2 DNA sequence generation and analysis

Total DNA was extracted from samples using a standard CTAB-chloroform/isoamyl alcohol method (Winnepennickx *et al.* 1993), and a fragment of the mtDNA cytochrome c oxidase subunit 1 (COI) gene was amplified by polymerase chain reaction (PCR) using universal pri-

mers(FishF1:5'-TCAACCAACCACAAAGACATTGGCAC-3';FishR1:5'-TAGAC TTCTGGGTGGCCAAA-GAATCA-3'; Ward *et al.,* 2005). PCRs were performed in a total volume of 20µl, containing 6µl of template DNA, 2 mmol/L $MgCl_2$, 0.5 µmol/L forward primer and 0.5 µmol/L of reverse primer, 0.2 mmol/L dNTP mix (20 µmol/L each dATP, dCTP, dGTP, dTTP), 1 × reaction buffer [75 mmol/L Tris-HCl,20 mmol/L (NH4)2SO4], and 0.1µl *Taq* polymerase (BioTaq, 5 U/µl). The PCR thermoprofile was 3 minutes at 95°C, followed by 35 cycles of 95°C for 30 seconds, 1 minute annealing at 54°C and 1 minute extension at 72°C, followed by a final extension stage for 3 minutes at 72°C. PCR products were then purified using SURECLEANPLUS (BIOLINE) and sequenced using both forward and reverse primers using BIG DYE technology. Sequences were aligned in BIOEDIT (Hall 1999) using the CLUSTALW programme (Thompson *et al.* 1994).

In order to add a greater depth of taxonomic and/or geographic coverage, COI sequences generated in the present study were supplemented with partial COI sequences downloaded from GENBANK on the 19th March 2018. Representative COI sequences (according the GENBANK labels) were available for all of the 28 Mugilidae species recognised as occurring across the Mediterranean, NE Atlantic, western African and South African coasts, except *Chelon bispinosus*. For all species (excluding *M. cephalus* owing to its broad geographic distribution and noted cryptic diversity (Livi *et al.* 2011) all available sequences were downloaded regardless of whether they were acquired from individuals sampled within the study region of the present work, though sequences lacking a location qualifier were excluded from analysis. Downloaded sequences were aligned with those generated in this study, any sequences that did not align and/or failed to cover a large proportion of the alignment were excluded from subsequent analyses.

The phylogenetic relationship among all sequences were inferred using Maximum likelihood trees constructed in MEGA v7 (Tamura *et al.* 2013) using the HKY +G +I model, which was identified as the optimum model based on the Akaike information criterion (Akaike 1974) implemented in ModelTest (AICc=16560.922, BIC=19960.765). Bootstrap values (BS) indicating level of support for each node were calculated after 1000 bootstrap replicates. Additional trees were constructed using Bayesian inference (BI), implemented in MrBayes 3.2 (Ronquist & Huelsenbeck 2003). Bayesian inference was also performed using the HKY+G+I model, with the assumption of unknown model parameters. Analysis was run using three heated and one cold chain over 10,000,000 generations, with the Markov chain sampled every 1000 generations and the first 15% of trees discarded as burn-in. Convergence was considered reached when the standard deviation of split frequencies fell below 0.01. Following Durand *et al.* (2012a,b) *Labracinus cyclophthalmus* and *Oryzias latipes* were used as outgroups as they were shown to be closely related to the Mugilidae in recent molecular investigations

of higher-level phylogenetic relationships in the Percomorpha (Chen *et al.*, 2007; Mabuchi *et al.*, 2007).

To investigate levels of within- and between-species/clade genetic differences, genetic distance was calculated in MEGA V7. Between and within species genetic distances and their standard error were calculated using the Tamura-Nei (Tamura & Nei 1993) substitution model (chosen as it was the optimum model that could be implemented in this analysis). To account for any bias in model selection, p distances were also calculated (data not presented, but comparable distances were resolved). Haplotypes were then partitioned based on their affinity to independent mtDNA lineages (clades), identified in the ML and BI topologies, and the range of genetic distances within and between putative clades (species) were recorded.

3.1.2.3 Timing of divergence within the Mugilidae phylogeny

Estimates of divergence times across all the major nodes of the Mugilidaee phylogeny were calculated using the RelTime (Tamura *et al.* 2012; 2013) method, implemented in MEGA v7. By relaxing the assumptions of a strict molecular clock, RelTime effectively estimates Timetrees from molecular sequence data when the evolutionary rate varies amongst lineages (Tamura *et al.* 2018; Battistuzzi *et al*. 2018). This method of divergence estimation has been shown to perform well in the analysis of large simulated (Tamura *et al.* 2018; Fillipski *et al.* 2014) and empirical (Mello *et al.* 2016) datasets, and has been shown to resolve divergence times similar to (Mello *et al.* 2016; Battistuzzi *et al*. 2018) or in some cases more accurate than (Sanderson 2003; Yang 2007), far more computationally expensive Bayesian methods (e.g. BEAST and MCMCTree).

To date the nodes of the Mugilidae phylogeny, RelTime was run using the HKY+G+I substitution model, and calibrated using 3 fossil calibrations (A, B & C), discussed further in Section 3.1.2.3.1. Through estimating branch specific and relative rates for each internal node (supported by the fact that when taxa are contemporaneous, the time elapsed from two sister linages most recent common ancestor will be equal), the RelTime method transforms the branch lengths (e.g. number of substitutions per site) of the starting phylogenetic tree into an ultrametric tree with branch lengths converted into relative time. Once this ultrametric tree is constructed, relative time is converted into absolute time (e.g. in calendar years), through the implementation of node calibrations, and 95% confidence intervals for these divergence time estimates can be calculated. The performance of Timetree estimates were assessed through comparisons of resolved node ages based on the Reltime method with those of previous studies using Bayesian methods (e.g. Santini *et al.* 2015).

3.1.2.3.1 Fossil calibrations

In order to convert the relative times calculated in RelTime into absolute time, 3 fossil calibrations were applied to the internal nodes of the Mugilidae phylogeny, following Santini *et al.* (2015):

Fossil calibration A was placed on the basal node separating the Mugil genera from all other Mugilidaee genera (e.g. Crenimugil, Osteomugil, Oedalechilius, Pseudomyxus, Parachelon & Chelon but see Neochelon), thus at the crown of the Mugil clade. This calibration point is derived from the fossilised remains of *Mugil princeps* (Switshenskaja 1973; Nolf 2003), the earliest recorded member of the Mugilidae. Accordingly the minimum and maximum boundaries applied to this calibration point were estimated at 54-30MYA (Santini *et al.* 2015).

Fossil calibration B was placed on the node separating the *Mugil cephalus* complex from its most recent common ancestor. The calibration of this node was based on the identification of otoliths, dated from the Miocene, with identical features as those of extant *M. cephalus* (Nolf & Aguilera 1998; Rey 1996; Grandstein *et al.* 2012). Based on this Santini et al. (2015) estimated the maximum and minimum boundaries of calibration point B at between 30-13.65MYA.

The final calibration point applied to the Mugilidae phylogeny was placed at the crown of the Chelon clade and was based on the identification of otoliths similar to those of *C. gibbosus* and *C. labrosus*, that were dated to the Oligo-Miocene (Reichenbacher & Weidmann 1992; de Leeuw *et al.* 2011), accordingly Santini *et al.* (2015) placed a minimum bound of 23MY and an upper bound of 32.2MY on this calibration point.

3.1.2.4 Assessment of species mis-assignment in publicly available databases.

Species and genus level misidentifications among the GENBANK sequences employed here were assessed based on ML and BI tree topology. Owing to the taxonomic complexity of the Mugilidae a reference database (Durand *et al.* 2012a,b) was chosen following Durand *et al.* (2017) to enable correct species assignment, with all clades/lineages resolved here given species/genus classifications according to the reference database. Where a COI sequence labelled as one species aligned with another species and/or genus (according to the reference database) a misidentification was recorded. Due to the longstanding ambiguity in the taxonomy and systematics of the Mugilidae, numerous cryptic species/lineages have been recognised within this family (Durand *et al.* 2015), but have not been appropriately characterised to date and as such lack morpho-anatomical species designations. Consequently misidentifications were only recorded when a sequence label failed to match a species within the reference data-base (Durand *et al.* 2012a) that is fully recognised taxonomically (e.g possesses genus and

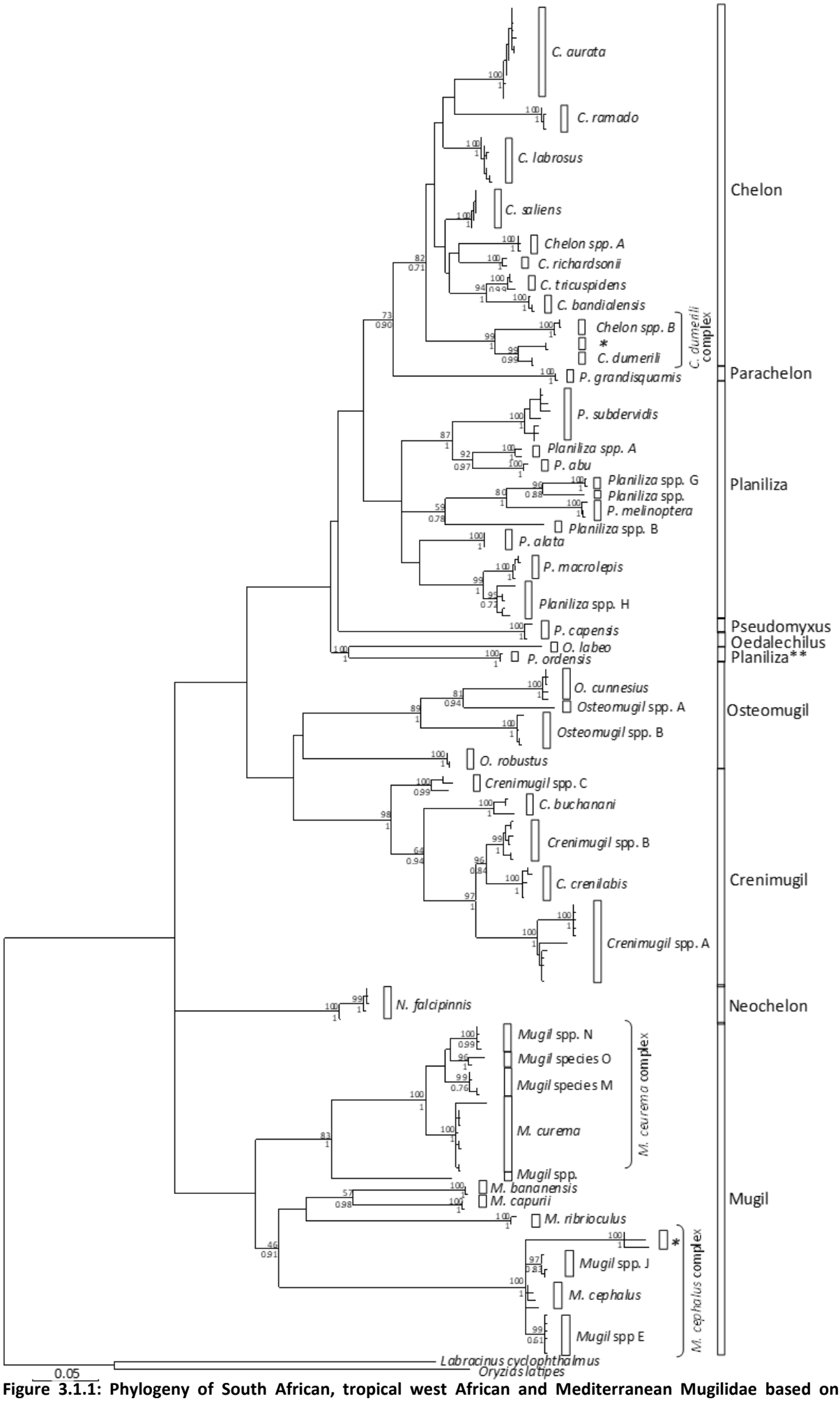


Figure 3.1.1: Phylogeny of South African, tropical west African and Mediterranean Mugilidae based on 516bp sequence of mtDNA COI. * indicates newly identified lineages. Maximum likelihood bootstrap and Bayesian inference values are given above and below branches, respectively. Species and genus nomenclature corresponds to reference database (Durand *et al.* 2012a)

species nomenclature, a type specimen and a morpho-anatomical description), although instances where bar-code miss-assignments fell within cryptic lineages/species (e.g. novel lineages identified by Durand *et al.* (2012a) but not yet fully characterised taxonomically) identified in the reference database, were also explored.

3.1.3 Results

After pruning, a 516bp fragment of the COI gene was aligned amongst 700 sequences. Samples collected by the present study generated 228 new sequences across all locations sampled (Turkey, N=15; Portugal N=13; Morocco N=6; Ghana N=19; Angola, Kwanza N=21; Angola, Flamingo N=23; SA Bergirivier N=21; SA Kneisner N=22; SA St. Francis N=7; SA Kariega N=20; SA Kowie N=8; SA Mpambanyoni N=53), and 472 were obtained from GENBANK. A total of 178 haplotypes were resolved across the 700 sequences, and assigned to 27 presumed species according to the Durand *et al.* (2012a) database.

3.1.3.1 Phylogenetic analysis of Mediterranean, west African and South African Mugilidae species.

As in previous molecular phylogenies (Durand *et al.* 2012a.b; Xia *et al.* 2016), the COI phylogeny (Fig. 3.1.1) reported here for grey mullet from the Mediterranean, west Africa and South Africa resolved 3 major clades corresponding to: 1) the Mugilinae, comprising the Mugil genus; 2) the Rhinomugilinae which contained the Osteomugil and Crenimugil genera; and 3) the Cheloninae which included Chelon, Parachelon, Planiliza, Pseudomyxus and Oedalechilus genera.

ML branching patterns (Fig. 3.1.1), and estimates of divergence dates (Fig. 3.1.2) indicate that the Mugilinae diverged first around 49.10MYA (95%CI=41.86-51.20), followed by a more recent split between the Cheloninae and Rhinomugilinae around 40.75MYA (95% CI=31.26-50.24). Across these major clades, the COI phylogeny (Fig. 1.1.1) resolved 9 haplogroups which correspond to previously designated genera (Chelon, Parachelon, Planiliza, Pseudomyxus, Oedalechilus, Osteomugil, Crenimugil, Neochelon and Mugil). Timetree estimates (Fig. 3.1.2) resolved an early origin of the Mugil genera (49.10MYA; 95%CI=41.86-51.20), followed by the diversification of Osteomugil and Crenimugil around 32.50MYA (95% CI=22.62MYA), with the 5 genera (Chelon, Parachelon, Planiliza, Pseudomyxus and Oedalechilus) that make up the Cheloninae the most recently derived (between 31.54-23.94MYA). The lowest between-genus genetic distances (Table 3.1.2) occurred between Chelon and Parachelon (0.128 to 0.178), with the highest values between Mugil and the other genera (0.186-0.410).

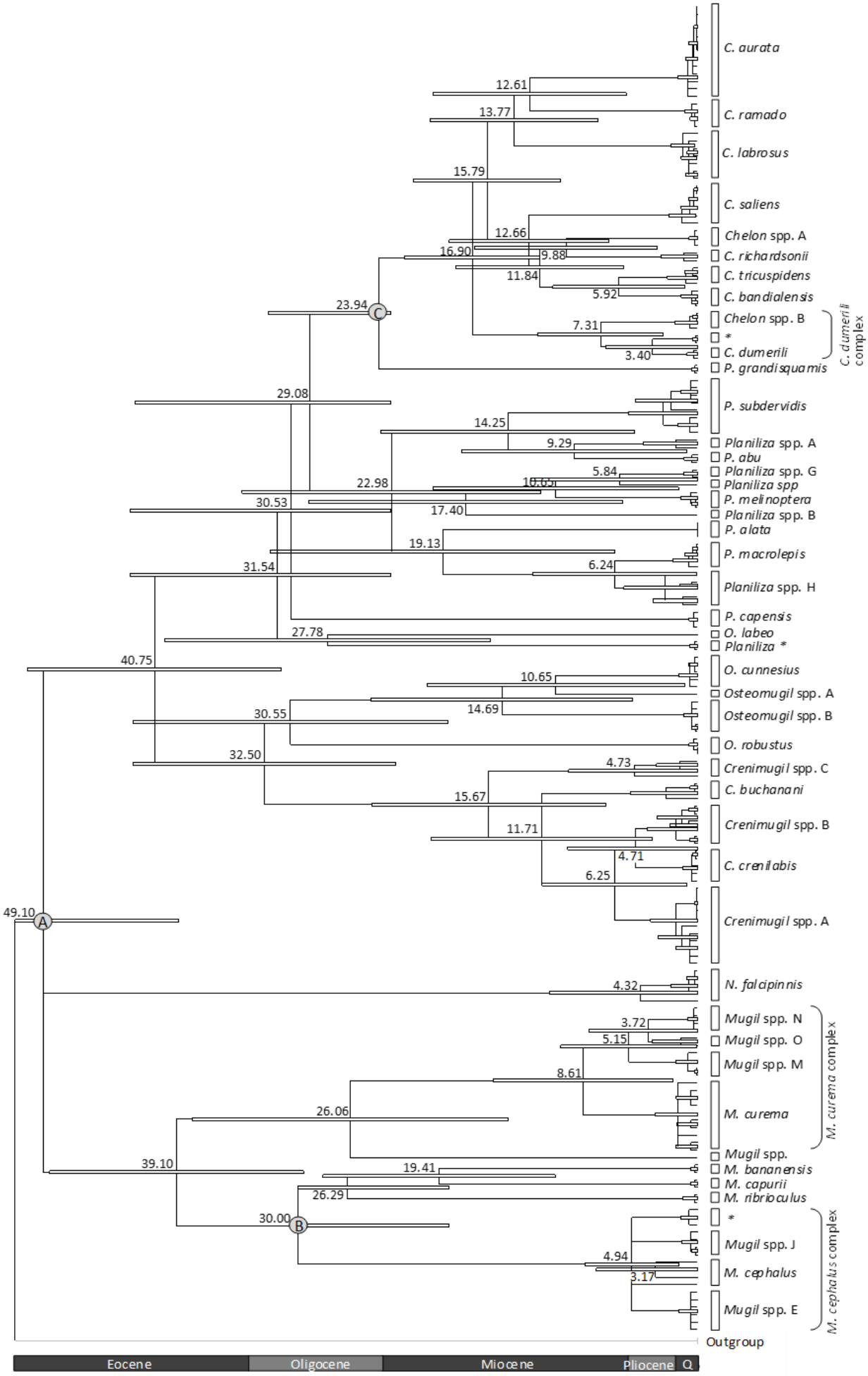


Figure 3.1.2: Chronogram showing estimated divergence times for the major Mugildaee lineages, calculated using the RelTime method applied to the ML topology generated from 516bp of COI. Calibrated nodes are indicated by a grey circle with letters corresponding to Fossil calibrations A, B and C in section 2.3.1. Estimated node ages are given on the node, with bars representing the 95% confidence interval of age estimates. Timescale, in geological epoch is provided on the x axis, with Q denoting the Quaternary.

Sequences labelled as *Planiliza alata* on GENBANK, and classified previously as belonging to the Planiliza genus (Durand *et al.* 2012a), formed a distinct and reciprocally monophyletic group labelled here as Planiliza*. The Planiliza* group clustered most closely with Oedalechilus, with genetic distances (Table 3.1.2.) between haplotypes of these two genera ranging from 0.163 to 0.166, although genetic distances between Planiliza* and confirmed Planiliza haplotypes were not substantially more divergent (TN93 = 0.133-0.235), with the majority of high genetic distance values associated with pairwise tests between *P. melinoptera* haplotypes.

Within the 10 putative genera identified here (including Planiliza*) a total of 47 reciprocally monophyletic and deeply diverged lineages were resolved in the COI phylogeny (Fig. 3.1.1). Which included all lineages previously identified within the genera/species included in the present investigation using a combination of nuclear and mitochondrial markers (Durand *et al.* 2012a,b, 2015; Xia *et al.* 2016), plus 2 new lineages (marked with * in Fig.3.1.1) not identified by the previous studies. Strong bootstrap (>96) and Bayesian inference (>0.76) support was resolved for all lineages, excluding Planiliza species B (BS=59, BI=0.78) which exhibited low bootstrap support and Mugil species E (BS=99, BI=0.61) where low Bayesian support was observed. Consequently all lineages/species are named according to the classifications of Durand *et al.* (2012b), and only findings that are novel from previous Mugilidae phylogenetic research (Durand *et al.* 2012a,b, 2017; Durand & Borsa 2015; Xia *et al.* 2016) are discussed

Table 3.1.2: Between- and within-genus Tamura-Nei (TN93) genetic distances across haplotypes of all Mugilidae genera included in the present analysis. * indicates a genus that does not correlate with previous phylogenies (see text).

	Chelon	Para-chelon	Planiliza	Pseudo-myxus	Oeda-lechilus	Planiliza*	Osteo-mugil	Creni-mugil	Neo-chelon	Mugil
Chelon	**0.030-**									
Parachelon	0.128-	**0.002**								
Planiliza	0.129-0.215	0.164-0.241	**0.039–0.210**							
Pseudomyxus	0.154-0.210	0.219-0.227	0.181-0.222	**0.002-0.008**						
Oedalechilus	0.181-0.251	0.226-0.226	0.190–0.224	0.167-0.173	**N/A**					
Planiliza*	0.152-	0.175-	0.133-	0.204-	0.163-	**0.002**				
Osteomugil	0.162–0.264	0.206-0.227	0.185–0.308	0.202–0.254	0.199-0.230	0.200-0.245	**0.130-0.190**			
Crenimugil	0.171–0.266	0.201-0.242	0.153-0.288	0.195-0.253	0.220-0.280	0.194-0.256	0.175–0.264	**0.041–0.167**		
Neochelon	0.197–0.251	0.244-0.252	0.187–0.253	0.222-0.257	0.237-0.257	0.218-0.231	0.194-0.253	0.185-0.287	**0.002-0.023**	
Mugil	0.211-0.410	0.233-0.373	0.192–0.398	0.220-0.366	0.216–0.315	0.254–0.342	0.200-0.345	0.202-0.415	0.186–0.335	**0.014-0.324**

further.

Although 15 divergent lineages/cryptic species have been identified and are currently recognised within the global *Mugil cephalus* complex (see Livi *et al.* 2011), the present analysis indicated a further lineage not previously identified, comprising GENBANK sequences from fish collected on the Mediterranean, Suez canal and Red Sea coasts of Egypt (KY652973.1 ; KY652973.1; KY780118.1). These 3 Egyptian sequences formed a monophyletic clade, highly genetically different (TN93= 0.063-0.102) from all other Mediterranean, west African and South African lineages of *M. cephalus*. In order to determine how these Egypt sequences were related to global populations of *M. cephalus*, COI sequences constituting all previously identified lineages/species of *M. cephalus* were compiled across a global distribution. When these global mullet sequences were included in phylogenetic analyses (Chapter 3.2: Fig. 3.2.5B) the Egypt samples clustered most closely with the western Atlantic lineage, which included samples from North and Central American populations of *M. cephalus* (Mugil spp. B in Durand *et al.* 2012a), however they still exhibited signals of a recent radiation from Mugil spp. B, forming a reciprocally monophyletic clade within Mugil spp. B.

3.1.3.2 Assessment of misidentifications and mislabelling within publicly available barcode (COI) data

3.1.3.2.1 Genus-level barcode/taxonomy incongruence

When assessing the phylogenetic analysis, a number of instances of mislabelling were identified among the publicly available COI sequences included in this study. Here genus level mislabelling is defined as the failure of a sequence to cluster phylogenetically within the genus it is labelled as belonging to (genus designations, and their recognised species, following Durand *et al.* 2012a), according to the species designation that it was listed under when obtained from GENBANK. In total 8 cases of genus-level misidentification were observed (Table 3.1.3). Within Chelon, Chelon species A was comprised solely of individuals from eastern

Table 3.1.3: Genus level mismatches identified in the GENBANK database with reassigned species/ genera based upon phylogenetic affinity to the reference data-base (Durand et al 2012a)

Accession number	GENBANK species label	Re-assignment
JF493804.1	*Planiliza melinoptera*	Chelon spp. A
JF493802.1	*Planiliza macrolepis*	Chelon spp. A
KU176385.1	*Planiliza macrolepis*	Chelon spp. A
JF494780.1	*Crenimugil seheli*	Osteomugil spp. B
JF494781.1	*Crenimugil seheli*	Osteomugil spp. B
KF489538.1	*Planiliza melinoptera*	Crenimugil spp. B
JQ060449.1	*Planliza alata*	Planiliza*
JQ060450.1	*Planliza alata*	Planiliza*

South Africa and Southern Mozambique, and accordingly included mis-labelled sequences obtained from South Africa and Mozambique labelled as *Planiliza melinoptera* and *Planiliza macrolepis* (JF493804.1 JF493802.1 KU176385.1). Within the cluster of species comprising the Osteomugil genus, GENBANK sequences from South Africa labelled as *Crenimugil seheli* (JF494780.1, JF494781.1) clustered with Osteomugil species B, like in Chelon species A, all sequences that made up Osteomugil species B were obtained from eastern South Africa and southern Mozambique. Within the Crenimugil genus, Crenimugil species B (a broadly Indo-Pacific clade/species) contained a sequence (KF489538.1), obtained from South Africa and labelled as *Planiliza melinoptera*. Two sequences obtained from Australia and labelled as *Planliza alata* (JQ060449.1 , JQ060450.1) fell outside of the cluster of species that make up the Planiliza genus, and were highly distinct from true *P. alata* (the distribution of which is thought to extend throughout the Indo-Pacific) that contained sequences solely obtained from the eastern coast of South Africa. These mislabelled *P. alata* sequences appeared to be more closely related to Oedalechilus but still formed a deeply diverged and reciprocally monophyletic lineage (Planiliza* - see above).

3.1.3.1.2 Within-genus species-level barcode/taxonomy misidentification

At the species level there were many more instances of apparent misidentification. When misidentifications were defined based on mismatches between a GENBANK sequence species designation and its phylogenetic affinity to recognised and cryptic species previously identified by Durand *et al.* (2012a), excluding the *M. curema*, *M. cephalus* and *C. dumerili* species complexes, there were a total of 70 (14.83%) cases of misidentification. However as many (if not all) of the cryptic species identified previously have no morpho-anatomical classification, any misidentification/mislabelling that occurred within these genetic clusters are excluded from further analysis and discussion. Consequently when only taking into account sequences that clustered within recognised species, 42 of the 472 (8.90%) sequences obtained from GENBANK aligned with a species that did not match the sequence label (Table 3.1.4), and when the genera level misidentification were included 50 of the 472 (10.59%) GENBANK sequence labels were mismatched. These misidentifications appeared to be more prevalent within particular genera, although all genera (excluding Crenimugil) with representative sequences from one or more species exhibited instances of misidentification.

Within Chelon a relatively low number of misidentifications were observed (N=5), which were all attributed to misidentifications between sympatric Mediterranean/north east Atlantic species: 1) *C. ramada* sequences being mislabelled as *C. saliens* (KX929966.1, KX929967.1, KX929964.1, KX929965.1); and 2) a *C. aurata* sequence (HM208834.1) being mislabelled as *C.*

Table 3.1.4: Species level mismatches identified in the GENBANK database with reassigned species/ genera based upon phylogenetic affinity to the reference data-base (Durand et al 2012a)

Genus	Accession number	GENBANK species label	Re-assignment
Chelon (N=5)	KX929966.1	*Chelon saliens*	*Chelon ramada*
	KX929967.1	*Chelon saliens*	*Chelon ramada*
	KX929964.1	*Chelon saliens*	*Chelon ramada*
	KX929965.1	*Chelon saliens*	*Chelon ramada*
	HM208834.1	*Chelon ramada*	*Chelon aurata*
Planiliza (N=31)	JQ060449.1	*Planiliza alata*	*Planiliza ordensis*
	JQ060450.1	*Planiliza alata*	*Planiliza ordensis*
	EF609544.1	*Planiliza macrolepis*	*Planiliza subviridis*
	EF609545.1	*Planiliza macrolepis*	*Planiliza subviridis*
	EF609546.1	*Planiliza macrolepis*	*Planiliza subviridis*
	EF609547.1	*Planiliza macrolepis*	*Planiliza subviridis*
	JQ060434.1	*Planiliza macrolepis*	*Planiliza subviridis*
	FJ347967.1	*Planiliza macrolepis*	*Planiliza subviridis*
	JQ045785.1	*Planiliza macrolepis*	*Planiliza subviridis*
	FN600159.1	*Planiliza carinata*	*Planiliza macrolepis*
	JQ623947.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500833.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500834.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500835.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500836.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500837.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500838.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500839.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500840.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500841.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500842.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500843.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500844.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500845.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500846.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500847.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500848.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500849.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500850.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500851.1	*Planiliza carinata*	*Planiliza macrolepis*
	KC500852.1	*Planiliza carinata*	*Planiliza macrolepis*
Mugil (N=6)	EU715465.1	*Mugil curema*	*Mugil rubrioculus*
	GU225396.1	*Mugil curema*	*Mugil rubrioculus*
	GU225397.1	*Mugil curema*	*Mugil rubrioculus*
	GU225398.1	*Mugil curema*	*Mugil rubrioculus*
	GU225399.1	*Mugil curema*	*Mugil rubrioculus*
	GU225400.1	*Mugil curema*	*Mugil rubrioculus*

ramada. Far more taxonomic confusion was observed within Planiliza, where 29 instances of incorrect species classification/mislabelling were observed, all of which occurred in areas where species were found in sympatry, these were attributed to: 1) sequences (JQ060449, JQ060450) that aligned in the tree with *P. ordensis* (a northern Australian species) but labelled as *P. alata*; 2) sequences labelled as *P. macrolepis* (EF609544.1, EF609545.1, EF609546.1, EF609547.1, JQ060434.1, FJ347967.1, JQ045785.1) clustering in the *P. subviridis* lineage; and 3) sequences labelled as *P. carinata* (FN600159.1, JQ623947.1, KC500833.1- KC500852.1) clus-

tering within the *P. macrolepis* lineage. Within Osteomugil, despite the inclusion of only one representative species of this genus, 2 cases of misidentification were observed, where GENBANK sequences labelled as *Crenimugil seheli* (JF494780.1, JF494781.1) clustered with Osteomugil species B (included in section3.1.2; Table 3.1.3). Finally, within the Mugil genus, and excluding individuals that fell within either the *M. cephalus* or *M. curema* species complexes, 6 instances of mislabelling were observed where GENBANK sequences obtained from the western Atlantic and labelled as *M. curema* (EU715465.1, , GU225396.1-GU225400.1) clustered with *M. rubrioculus* (a western Atlantic and Caribbean species). No instances of misidentification were observed within genera for which only one representative species was included in analysis (Oedalechilus, Parachelon, Neochelon Pseudomyxus; but see Osteomugil). Additionally no mislabelling/misidentifications where observed within the Crenimugil genus.

3.1.4 Discussion

3.1.4.1 Assessment of misidentifications and mislabelling within publicly available barcode (COI) data.

A striking finding of this study was the frequency of mismatches between species designations based on mtDNA lineage and those based on species labels from publicly available sequence databases, with 10.59% of sequences harvested from GENBANK clustering phylogenetically with haplotypes from another species or even genus. Similarly high proportions of misidentifications within Mugilidae genetic datasets have been reported previously (Durand *et al.* 2017).

One explanation for these mismatches between assigned species designations and their mitochondrial lineage could be mitochondrial introgression, due to past hybridisation between closely related species, followed by subsequent backcrosses (Freyhoff *et al.* 2005). Although mitochondrial introgression has been cited to explain paraphyly in the phylogenies of many species (reviewed in Funk & Omland 2003; in teleosts by Avise & Saunders 1984), we lack the comparative nuclear genetic data required to test this hypothesis in the present dataset. However as hybridisation drives introgression, here it would be expected that under an introgression scenario instances of mtDNA lineage-species label discordance should occur solely within closely related and/or sympatric lineages. Conversely here we observe many instances of sequences that aligned phylogenetically with a different genus or a species highly divergent from the species designation assigned to it on GENBANK. Consequently it is most likely that the majority of mismatches between species labels and mtDNA lineage are either an artefact of contamination and/or human error in laboratory/analytical procedures, or more likely are representative of true misidentifications of species prior to molecular genetic analysis.

The frequency of misidentifications observed here and in previous literature (Durand *et al.* 2017) are likely attributable to the morphological similarity among Mugilidae species and genera. The correct species, and often genus, designation of grey mullets is reliant on a comprehensive taxonomic and systematic understanding of this family, as distinguishing characters are complex to discern, with available identification keys either failing to include all species (Albaret 1984; 1987; Albaret & Legendre 1985; Séret and Opic, 1986), or containing overlapping diagnostic features for species (Albaret 2003; Harrison 2007). The lack of a comprehensive and easily accessible means of identification for many Mugilidae species and genera may also explain the higher frequency of misidentification in species-rich and taxonomically poorly defined genera, such as Planiliza, when compared with genera such as Mugil which comprise more species but are arguably represented and defined in greater detail in taxonomic and systematic literature (Albaret 2003; Harrison 2007).

Durand *et al.* (2017) noted that in molecular investigations of the Mugilidae, misidentifications were more frequent in barcoding studies when compared with phylogenetic and phylogeographic investigations, and attributed this to the reliance of barcoding studies on morphological classification in the place of more extensive phylogenetic/phylogeographic assessments. Here we contribute to this argument by identifying numerous voucher specimens that do not conform phylogenetically to their morpho-anatomical species designations. This prevalence of misidentifications in barcoding databases indicates that much of the species richness and abundance of the Mugilidae is misclassified, and as such the failure of such studies to accurately identify grey mullets to the species or genus level has serious implications for not only the scientific investigation of this family, but also the successful management and conservation of an important fisheries resource. Accordingly an important finding of this study was that the single locus COI phylogeny resolved all expected lineages based on the reference database (Durand *et al.* 2012a) and demonstrated similar strength to resolve species level divergence as studies incorporating a suite of mtDNA and nuclear markers (Durand *et al.* 2012a,b; Durand & Borsa 2015; Xia *et al.* 2016). As such we highlight the utility of the COI barcoding locus as a relatively cost effective means of facilitating species-specific monitoring and management practices for the Mugilidae family.

3.1.4.2 Phylogenetic analysis of Mediterranean, west African and South African Mugilidae species.

The mtDNA phylogeny reported here was characterised by relatively short internal branches, whereas terminal branches were generally longer, a branching pattern that has been attributed to a rapid radiation during the initial diversification of the Mugilidae (Durand *et al.* 2012a). In accordance with this, all divergence date estimates for genera level splits in the phylogeny were between 49.10MYA and 23.94MYA. The phylogenetic tree (Fig. 3.1.1) resolved here was

analogous to previous molecular phylogenies of this family (Durand *et al.* 2012a,b; Durand & Borsa 2015; Xia *et al.* 2016), and consequently only those results that are novel, or contrast with previous findings, will be discussed in greater detail.

According to Xia *et al.* (2016) Neochelon should occupy a basal position within the Cheloninae, however in Fig. 3.1.1 and in Durand *et al.* (2012a;b) this genus takes the position of a fourth major clade between the Mugilinae and the Cheloninae/Rhinomugilinae. However the phylogenetic position of Neochelon was supported by low BS and BI values and was based on only a single mtDNA marker, unlike the combined mtDNA and nuclear genetic analysis of Xia *et al.* (2016).

A major contrast to previous molecular phylogenies (Durand *et al.* 2012a,b; Durand & Borsa 2015; Xia *et al.* 2016) was the phylogenetic placement of a second lineage of *Planiliza alata* (lablelled as Planiliza* in Fig. 3.1.1 & 3.1.2, and referred to as *Planiliza ordensis* here after Durand *et al.* 2015), which clustered more closely with the Oedalechilius genus than other Planiliza. The paraphyly of *P. alata* has been reported previously (Durand *et al.* 2012a), with true *P. alata* (based upon the Madagascan location of the type specimen) occurring in samples from the eastern coast of Africa and the second lineage (labelled here as *P. ordensis*) occurring in samples from northern Australia. Unlike many other cryptic grey mullet lineages the differentiation of the east African and Australian lineages of *P. alata* has been confirmed morphologically (Larson *et al.* 2013), which led Durand *et al.* (2015) to call for the revival of the species name *P. ordensis* (Thomson 1997) to represent the distinct Australian clade. However, in contrast to the previous phylogenetic investigations in which both P. alata lineages were included (Durand *et al.* 2012a, Xia *et al.* 2016), here *P. ordensis* failed to cluster within the Planiliza clade. The failure of P. ordensis to cluster within the Planiliza genus raises concerns as to the current systematics of the Planiliza genus, primarily as *P. ordensis* is the type specimen of this genus (Whitley 1945), and the phylogenetic positioning of *P. ordensis* in previous studies (Durand *et al.* 2012a) influenced the proposed resurrection of the Planiliza genus (Durand *et al.* 2012b). However, genetic distances reported here between haplotypes belonging to *P. ordensis* individuals and those that made up the Oedalechilus genus, were of a similar magnitude to those between *P. ordensis* and the Planiliza genus. It is therefore more likely the case that the short section of COI implemented here failed to fully resolve the intergeneric relationships of *P. ordensis*, particularly when compared with those investigations that employed more mtDNA (Durand *et al.* 2012a) and nuclear markers (Xia *et al.* 2016). Consequently this finding and its potential implications to Mugiliade taxonomy should be viewed tentatively.

Through the inclusion of novel sequences generated in this investigation, alongside sequenc-

es obtained from GENBANK which (to the best of our knowledge) have not been incorporated into any previous phylogenetic investigation of grey mullets, this study resolved two previously unidentified lineages of grey mullets. The first was a unique Angolan lineage of the *Chelon dumerili* species complex (discussed further in chapter 3.2). The second was represented by three sequences of *Mugil cephalus* obtained from GENBANK and labelled as being sampled from all three basins that make up the Egyptian coast (Red sea, Mediterranean Sea and Suez Canal). These sequences fell within a clade which was highly divergent from, but still shared a most recent common ancestor with, all other Mediterranean, west African and South African *M. cephalus* sequences, with estimated divergence dates placing this clades divergence within the rapid diversification of all major *M. cephalus* clades around 4.94MYA. When these sequences were included in a phylogeny representing the global distribution of *M. cephalus* (Chapter 3.2, Fig. 3.2.5B), they clustered with a lineage thought to be restricted to eastern America and Mexico (Mugil species B: Durand *et al.* 2012a), but were still highly differentiated from this American lineage, suggesting a signal of a recent radiation.

Mugil cephalus is an important species in the aquaculture industry of Egypt, and although natural populations have been farmed in traditional ways for centuries, more recently Egyptian aquaculture has moved to more intensive practices (El-Sayed 2015). Although to the best of our knowledge there are no reports of translocation of *M. cephalus* stocks for the purpose of Egyptian aquaculture, introductions and translocations of *M. cephalus* for the purpose of aquaculture have been noted in Turkey (Innal & Erk'akan 2006), as well as translocations from Israel into Greek mullet farms (Economidis *et al.* 2000). It is therefore not implausible that the occurrence of this geographically distant western Atlantic lineage within Egyptian waters may be the product of this intensification of grey mullet aquaculture, where Egyptian farms have been restocked with mullets of western Atlantic origins. The possible introduction of non-native fish is supported by evidence of genetically distinct 'wild' and 'aquaculture' populations within the Red Sea (Magdy *et al.* 2016). Additionally, sequences from Egypt showed a degree of separation from the eastern Atlantic *M. cephalus* clade, forming a reciprocally monophyletic lineage, which could be the genetic signature of a recent bottleneck and divergence, encouraged by the strong selection pressures and genetic admixture placed upon aquaculture species. Without knowing conclusively the purpose and source of the original Egyptian sequences these results may simply be attributable to imprecise labelling of sequences within the reference database, or sequencing error and/or contamination when generating these sequences.

3.1.4.3 Conclusion

Overall these results have implications for the management of Mediterranean, west African and South African grey mullets. The observed levels of cryptic diversity across the study region highlights the potential for inaccurately managed harvesting, with previously uncharacterised Mugilidae lineages identified in Angola (*C. dumerili*) and Egypt (*M. cephalus*). A high proportion of cryptic diversity was identified in South African waters, with sequences obtained from individuals collected along the eastern South African coast aligning with four previously characterised (Durand & Borsa 2015) putative cryptic species (Chelon species A, Osteomugil species B, Crenimugil species A, Crenimugil species B). The South African cryptic species were all sampled from the tropical/subtropical KwaZulu-Natal coast of eastern South Africa, and based on the location qualifier given to their comprising sequences, appear to constitute either eastern South African endemic (Osteomugil species B), SW Indian Ocean-restricted (Chelon species A) or broadly Indo-Pacific (Crenimugil spcies A & B) lineages. The high proportion of cryptic diversity within the southwestern Indian Ocean samples likely reflects the species richness of the tropical Indo-Pacific, a recognised biodiversity hotspot (Ridgway & Sampayo 2007), in addition to the unique oceanographic features of the region, that have been shown to promote eco-evolutionary diversification and cryptic speciation in a range of taxa (Thorpe *et al.* 2000; Borsa *et al.* 2013; Healey *et al.* 2018a,b Chapters 5.1 & 5.2). Again this cryptic diversity highlights the risks of indiscriminate harvesting of Mugilidae, particularly along the tropical coasts and estuaries of eastern South Africa, as well as the potential for inaccurate estimations of statistics essential for appropriate management, such as species richness and estimations of biological, ecological and evolutionary processes (Garcia-Vazquez *et al.* 2012).

The cryptic diversity revealed within grey mullets, in combination with the frequency of misidentifications observed within publicly available databases, highlights the potential for a lack of accurate species-level resolution in management initiatives and fisheries statistics, and could result in an over- and/or under-representation of stock abundance (Griffiths & Heemstra 1995). Of the 28 Mugilidae species recognised as occurring in the Mediterranean, west coast of Africa and South Africa, only 8 possess any species-specific fisheries statistics (FAO data accessed 19/04/2018). Of those species with fisheries assessments a clear geographic bias is apparent, with the majority of Mediterranean species (6 of 8) possessing species-specific fisheries statistics. Conversely along the west and southern African coasts the only species that have been assessed are those with global or inter-oceanic distributions (*M. cephalus* and *M. curema*), Mediterranean species whose ranges extend into tropical west Africa (*C. labrosus*), or species that are of particular commercial importance in Saudi Arabia and Iran

(*C. seheli*.) As such the validation of COI as a cost effective and informative marker for the identification of Mugildae to the species level, could contribute to the formulation of more species-specific and refined management practices for the Mugilidae. These results are therefore important, not only for the purpose of sustainable fisheries, but also as many Mugilidae species play pivotal roles in the functional ecology of coastal ecosystems globally (Lebreton *et al.* 2011, Chang & Iizuka 2012, Whitfield *et al.* 2012). Ultimately these results highlight the need for a comprehensive revision of Mugilidae fishery practices and conservation, based on a revision of the taxonomy and identification of all Mugilidae species, and highlights the utility of the universal COI barcoding locus as an effective means of facilitating this through enabling accurate species identification.

CHAPTER 3.2

Comparative phylogeography of grey mullets across the Atlantic-Mediterranean boundary, the tropical coast of west Africa and around the Benguela

3.2.1 Introduction

The comparison of phylogeographic structure amongst multiple co-distributed taxa (comparative phylogeography) represents a powerful tool in delineating the relative influence of environmental stochasticity versus ecological determinism in shaping the evolutionary diversification of the biota of a region (Barber *et al.* 2006; Hickerson & Meyer 2008; Dawson 2012; 2014; Bowen *et al.* 2016). The grey mullets (Mugilidae), a species-rich fish family consisting of around 73 recognised species (Eschmeyer & Friche 2015), provide an ideal opportunity to test such concepts of eco-evolutionary diversification. The Mugilidae occupy inshore and coastal habitats in tropical, subtropical and temperate regions (Thomson 1966), and often utilise estuarine and near-freshwater environments for at least some of their life cycle. Across western African and Mediterranean coasts an array of mullet species occur, with groups of species sharing similar distributional patterns that can at least superficially be associated with recognised biogeographic boundaries and oceanographic features. However some grey mullet species, such as *Mugil cephalus*, have far broader geographical distributions that appear to be unconstrained by recognised biogeographic boundaries (Briggs 1960). This apparent discordance in species distributions raises questions as to the universality of drivers of speciation, distribution and gene flow across species in this morphologically and ecologically conservative family.

Within the Mediterranean Sea eight species of Mugilidae are currently recognised: *M. cephalus, Chelon ramado, Chelon labrosus, Oedachelius labeo, Chelon aurata, Chelon saliens, Planiliza abu* and *Planiliza carinata* (see Table 3.2.1; Thomson 1997). With the exception of *P. carinata and P. abu*, which are restricted to the periphery of the eastern Mediterranean, the distribution of these species extends into the eastern Atlantic. For these Mediterranean-Atlantic species, the southern limit to their distribution appears to extend as far south as the Atlantic coasts of Morocco (*C. saliens*, *O. labeo* and *C. aurata*), Mauritania (*C. ramado*) and Senegal (*C. labrosus*), whereas *M. cephalus* has a global distribution that within the eastern Atlantic extends as far as South Africa.

Along the tropical western African coast (defined here as Senegal to Angola) 10 species of grey mullets (see Table 3.2.1) are currently recognised (Thomson 1997, Albaret *et al.* 2003; Trape & Durand 2011; Trape *et al.* 2012), with *Chelon bandialensis*, *Mugil bananensis*, *Neo-*

chelon falcipinnis, Parachelon grandisquamis, *Mugil capurii* and *Chelon bispinosus* restricted to this region. Three species (*M. bananensis*, *N. falcipinnis* and *P. grandisquamis*) have ranges that extend throughout the west African tropics, whereas the ranges of others are restricted to localised island populations (*C. bispinosus*) or the Senegal - Guinea-Bissau region (*M. capurii* and *C. bandialensis*). The remaining four grey mullet species found in tropical west Africa are either restricted to a more southern distribution for which Angola constitutes their northern limit (*C. tricuspidens*) or are species for which the west African region is part of a more extensive inter-oceanic distribution (*M. cephalus*, *Mugil curema and Chelon dumerili*).

Table 3.2.1: Mugilidae species recorded from the four biogeographic regions considered in the present study, indicating where species are restricted to single regions and where species occur across multiple regions.

Mediterranean	NE Atlantic	Tropical west Africa	South Africa (temperate)	South Africa (tropical)
P. carinata		*N. falcipinnis*		*O. cunnesius*
P abu		*M. capurii*		*O. robustus*
		M. curema		*C. buchanani*
		M. bananensis		*C. seheli*
O. labeo				*P. alata*
				P. macrolepis
				P. subdervidis
				P. melinoptera
				C. crenilabis
C. aurata		*P. grandisquamis*	*P. capensis*	
C. ramado		*C. bispinosus*	*C. richardsoni*	
C. saliens		*C. bandialensis*	*C. tricuspidens*	
C. labrosus		*C. dumerili*		
M. cephalus				

The temperate and sub-tropical waters of South African estuaries and coastal waters harbour a number of Mugildae, with 14 species (see Table 3.2.1) recorded as occurring across the western, southern and eastern coasts (Whitfield 1998; Harrison 2003; Heemstra & Heemstra 2004). South African mullet species can be subdivided into groups based on their optimal environmental conditions (James *et al.* 2016; Harrison & Whitfield 2006). The first and most species-rich group includes tropical Indian Ocean and Indo-Pacific species (*Osteomugil cunnesius, Osteomugil robustus, Crenimugil bunchananni, Crenimugil seheli, Planiliza alata, Planiliza macrolepis, Planiliza subdervidis, Planiliza melinoptera* and *Crenimugil crenilabis*) whose westerly distribution extends as far as the tropical and sub-tropical waters of Kwazulu-Natal and the northern part of the Eastern Cape. Although all other South African mullet species (*C. dumerili, Chelon tricuspidens, Chelon richardsonii, M.cephalus* and *Pseudomyxus capensis*) have been recorded as occurring in tropical Indian Ocean waters, at least as far north as Kwazulu-Natal, their optimum environmental niche appears to lie in cooler waters further south. Within the warm temperate waters of the Eastern Cape the most abundant species are *C. dumerili*, *C. tricuspidens* and *P. capensis*, whereas in the cool-temperate waters of the western South African coast the South African endemic *C. richardsonii* is the most abundant species. The distributions

of *M. cephalus,* C*. dumerili* and *C. tricuspidens* also extend along the west coast of South Africa into the cold waters of Namibia and into southern Angola.

Although many Mugilidae species appear to have shared distributions that can loosely be associated with recognised biogeographic boundaries across the Mediterranean, the west coast of Africa and South Africa, it remains unclear as to the actual environmental and eco-evolutionary drivers of these patterns. This is particularly so for species (e.g. *M. cephalus* and *C. dumerili*) where these biogeographic boundaries appear to have had little influence on their extensive, often inter-oceanic distributions. The aim of the present study was to investigate relationships of known biogeographic provinces and oceanographic features with patterns of genetic diversity, gene flow and speciation events in the Mugilidae across the region from the Mediterranean to South Africa. In order to do this, datasets were compiled for species which had mtDNA COI sequences (either generated in this study, or compiled from GENBANK) representative of locations across these biogeographic boundaries/regions. In total 5 species fit these criteria across the Mediterranean-Atlantic region (*C. aurata, C. saliens, C. ramado, C. labrosus* and *M. cephalus*), with 5 species spanning the tropical west African coast (*M. cephalus*, M*. curema, M. bananensis, N. falcipinnis* and *C. dumerili*), and 3 species represented between Angola and South Africa spanning the BUS (*C. tricuspidens, C. dumerili* and *M. cephalus*). Species-specific patterns of gene flow and their association with historical vicariance were described for each biogeographic boundary, alongside any similarities in evolutionary subdivision (and timescales) between species and species groups that may be indicative of a shared biogeographic history.

3.2.2 Methods

3.2.2.1 Sample acquisition and DNA sequence generation

As set out in Chapter 3.1, in order to cover the distributions of the target species for this thesis grey mullet samples were collected from commercial and artisanal fish markets, as well as targeted fishing, at 11 sites spanning the Mediterranean, NE Atlantic, west Africa and South Africa. Total DNA was extracted from each fin clip tissue sample, and used to amplify the COI mtDNA gene region as described in Chapter 3.1.

In order to add a greater depth of taxonomic and/or geographic coverage, sequences generated in this investigation were supplemented with those publicly available on GENBANK. Partial COI sequences were downloaded from GENBANK on the 19th March 2018. Based on recognised distribution patterns, 12 mullet species were identified as candidates to test patterns of diversification across the Mediterranean-Atlantic boundary region, across the tropical region of west Africa, and across the BUS. Of these 12, 10 species (*C. saliens, C. labrosus, C.*

Table 3.2.2: Locations and number of samples of five grey mullet species used in comparative phylogeographic analyses across the Atlantic-Mediterranean transition. * indicates species whose distribution extends across multiple regions (see Table 3.2.1).

		C. aurata	C. saliens	C. ramado	C. labrosus	M. cephalus*
Mediterranean	**Turkey: general**	21				23
	Turkey: Istanbul	2	1	1		
	Turkey: Fethiye	13	2	1	1	2
	Turkey: Iskenderun		1	1	1	
	Greece: Mediterranean	6	3		7	3
	Italy	1	1	2	1	
	Spain: Palamos	3		1	3	
	France: Martigues					1
	Tunisia: La Goullette	1	2	1	1	1
Atlantic	**Portugal general**			5		
	Portugal: Olhao				13	
	Morocco Atlantic		2	1	2	
	Morocco: Merja Zerga			1		1
	Morocco: Oeud Iqem			3	2	1
	Morocco: Essaouira	6				
	Morocco: Agadir	2			1	1
	Morocco: Khnifniss	2			1	2
	Western Sahara: Dhakla	1				

Table 3.2.3: Locations and number of samples of 5 grey mullet species used in comparative phylogeographic analyses across tropical west Africa. * indicates species whose distribution extends across multiple regions (see Table3.2.1). Numbers in brackets represents the number of samples amplified for the nuclear alpha-amylase-intron 1 locus.

	M. curema	M. bananensis	N. falcipinnis	C. dumerili*	M. cephalus*
Senegal: general					1
Senegal: Saint Louis			1	1	
Senegal: Dakar					2
Senegal: Saloum	2	1	1		
Senegal: Toubacouta			1	1	
Guinea-Bissau				1	
Guinea					1
Ivory Coast		3			
Ghana: Tema	6		13		
Togo: Lome	2		2	1	2
Benin	2				1
Gulf of Guinea: general		1	1	2	
Angola: Kwanza	1	1		9 (8)	9
Angola: Cunene river					1
Angola: Flamingo lodge					3

ramado, C. aurata, M. bananensis, N. falcipinnis, M. curema, C. tricuspidens, C. dumerili & *M. cephalus*) had representative sequences available (generated in this study or obtained from GENBANK) from both sides of the proposed boundaries, and consequently were included in further analyses (Tables 3.2.2, 3.2.3 & 3.2.4). The downloaded sequences were aligned with those generated in this study, and any sequences that did not align and/or failed to cover a large proportion of the alignment were excluded. Potential species mis-identification and mis-assignment as a result of inconsistencies in GENBANK records were eliminated during the proof testing carried out in Chapter 3.1.

Table 3.2.4: Locations and number samples of 3 grey mullet species used in comparative phylogeographic analyses across the Benguela upwelling system. * indicates species whose distribution extends across multiple regions (see TableX.1). Numbers in brackets represents the number of samples amplified for the nuclear alpha-amylase-intron 1 locus.

		C. tricuspidens	C. dumerili*	M. cephalus*
North Benguela	**Angola: Kwanza**		9 (8)	9
	Angola: Cunene river			1
	Angola: Flamingo lodge	20		3
South Benguela	**South Africa: Sundays river estuary**	1		
	South Africa: Kariega river		19 (5)	1
	South Africa: Kowie river estuary		24 (7)	
	South Africa: Great fish river		2	2
	South Africa: KZN general			3
	South Africa: Mpambanyoni river		24	12
	South Africa: Inyoni rocks	1		
	South Africa: Umhlanga		3	
	South Africa: St Lucia			3

3.2.2.2 Phylogeographic analysis

Alignments containing 509bp of the COI gene were compiled for each species. Median-joining haplotype networks were constructed using NETWORK 5.0 (Fluxus Technology). Genetic distances between haplotypes were calculated using the Tamura-Nei substitution model and p distances, implemented in MEGA V7 (Tamura *et al.* 2013), and the range of within- and between-clade distances were explored. Owing to the variability in sample sizes and geographic coverage of samples it was decided that any additional calculation of within location/clade diversity, neutrality or genetic differentiation would be too heavily influenced by sampling regime. However for *C. dumerili* and *M. cephalus* a more comprehensive sample set allowed for the construction of maximum likelihood (ML) topologies. ML trees were constructed in MEGA v7, using the HKY +G +I model, which was identified as the best-fit model based on AIC values (Akaike 1974), implemented in ModelTest, with bootstrap (BS) support calculated after 1000 replicates. Finally the age of intra- and inter- specific divergence events were estimated using the RelTime (Tamura et al 2012; 2013) method, implemented in MEGA, with node ages estimated based on 3 fossil calibrations applied to the entire Mugildae ML phylogeny (see Chapter 3.1; section 2.3 for a comprehensive description of the analysis applied to this dataset).

Due to the identification of a previously uncharacterised lineage within the *C. dumerili* dataset a region of the nuclear alpha-amylase intron 1 locus was amplified to assess nuclear genetic patterns across *C. dumerili* clades/species. PCRs used primers described in Hassan *et al.* (2002): Am2B1F 3' CCTTCATCTTCCAGGAGGTAC'5; Am2B1R 3' TTCACCTCCCAGATCAATAAC'5. PCRs were performed in a total volume of 20µl, containing 6µl of template DNA, 2 mmol/L $MgCl_2$, 0.5 µmol/L forward primer and 0.5 µmol/L of reverse primer, 0.2 mmol/L dNTP mix (20 µmol/L each dATP, dCTP, dGTP, dTTP), 1 × reaction buffer [75 mmol/L Tris-HCl,20 mmol/L

$(NH_4)_2SO_4$], and 0.2U *Taq* polymerase (BioTaq, 5 U/µl). The PCR thermo-profile was: 5 minutes at 94ºC, followed by 35 cycles of 94ºC for 30 seconds, 45 seconds annealing at 55ºC and 45 seconds extension at 72ºC, followed by a final extension for 7 minutes at 72ºC. Amplicons were purified using SURECLEANPLUS (BIOLINE) and sequenced in both directions using BIG DYE technology. A 643bp fragment of alpha amylase intron 1 was aligned across samples. As no heterozygote base positions were detected subsequent alignment and analysis was identical to those described above for COI.

3.2.3 Results

3.2.3.1 Phylogeographic and speciation patterns among Mugilidae species groups across the Mediterranean - NE Atlantic, tropical west African and southern African biogeographic regions

In total a 509bp fragment of the COI gene was amplified for 323 individuals across 10 Mugilidae species. When inspecting the ML phylogenies resolved across the entire Mugilidae COI alignment (Chapter 3.1. Fig. 3.1.1 & 3.1.2) a clear association with patterns of inter-specific phylogenetic partitioning and the contemporary geographic distribution of species was apparent, particularly within the Chelon genus (Fig. 3.2.1). The species that make up the Chelon genus can be associated with 5 geographic distributions. *C. aurata, C. ramado, C. labrosus and C. saliens* all have distributions that extend throughout the Mediterranean basin and the north east Atlantic, as far south as the north-west African coast. Both *C. dumerili* and *C. bandialensis* are species that appear to be restricted to the west African tropics, occurring from the Gulf of Guinea, north to Guinea-Bissau and Senegal. Two more groups of species can be found either restricted to the northern (the Angolan clade within the *C. dumerili* complex) or southern (*C. richardsonii*, Chelon spp. B) Benguela subsystems or with representatives in both the northern and southern BUS, as is the case for *C. tricuspidens*. Finally Chelon spp. A appears to be restricted to the subtropical/ tropical Indian Ocean waters of Mozambique and the north-west South African coast.

The Chelon genus can be further subdivided into 3 clades, the first of which diverged from the rest of the members of the genus around ~16.9 MYA (95% = 21.97-11.82) and comprised the species of the *C. dumerili* complex, which subsequently diverged between tropical west Africa and South Africa ~7.3 MYA, and finally between the Gulf of Guinea and Angola ~3.4 MYA. The second clade comprised solely of Mediterranean and north east Atlantic species (*C. aurata, C. ramado* and *C. labrosus*) and diverged from the third clade, which consists of species distributed across the Mediterranean, eastern Atlantic and Indian oceans (*C. saliens*, Chelon spp. A*, C. richardsonii, C. tricuspidens, C. bandialensis*) around 15.79 MYA (95% CI = 21.31-10.27). Within this third clade an initial divergence of Mediterranean-north east Atlantic species (*C.*

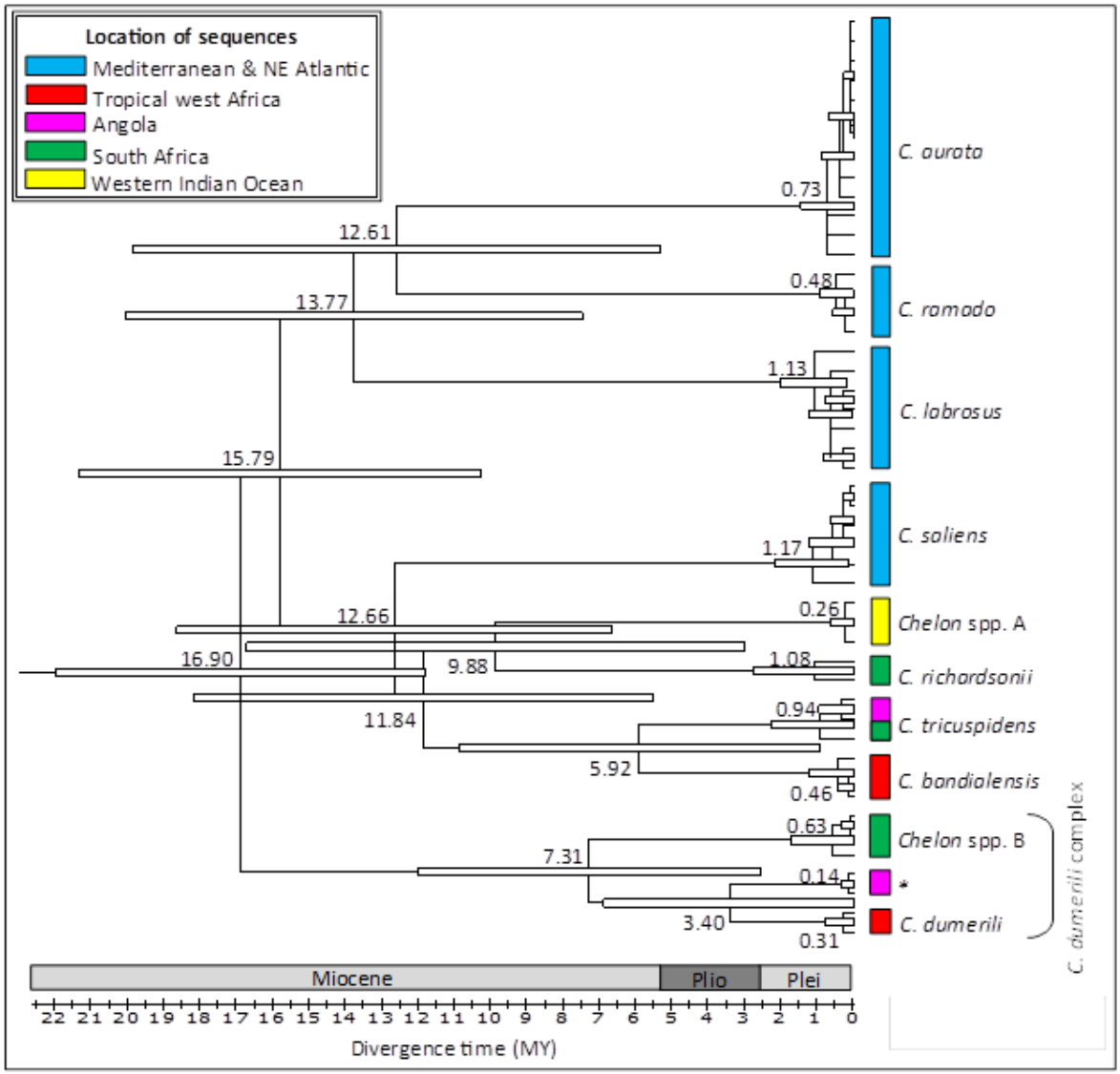


Figure 3.2.1: Chronogram showing estimated divergence times within the Chelon genus, calculated using the RelTime method applied to the ML topology generated across all Mugilidae species using 516bp of COI (see Chapter 3.1). Estimated node ages are given on the node, with bars representing the 95% confidence interval of age estimates. Timescale, in million years (MY) alongside the geological epoch is provided on the x axis, with Plio denoting the Pliocene and Plei denoting the Pleistocene. Coloured bars represent the contemporary distribution of the indicated species. * represents a novel lineage that has been resolved for the first time in this investigation.

saliens) from those distributed across the rest of the tropical west African and southern African coasts was estimated to have occurred around 12.66 MYA (95% CI = 18.65-6.66), followed by the bifurcation of south African (*C. richardsonii*) and Indian Ocean (Chelon spp. A) species from tropical west African (*C. bandialensis*) and northern and southern Benguelan (*C. tricuspidens*) species around 11.84MYA (95% CI = 18.16-5.52). Following this, around 9.88 MYA (95% CI = 16.75-3.02) the Indian ocean Chelon spp A. and South African *C. richardonii* diverged from their most recent common ancestor, and finally around 5.92 MYA (95% CI = 10.88-0.97) tropical west African *C. bandialensis* diverged from *C. tricuspidens*. All within-species diversification within the Chelon genera are estimated to have occurred relatively recently, since the mid Pleistocene (1.14-0.14 MYA).

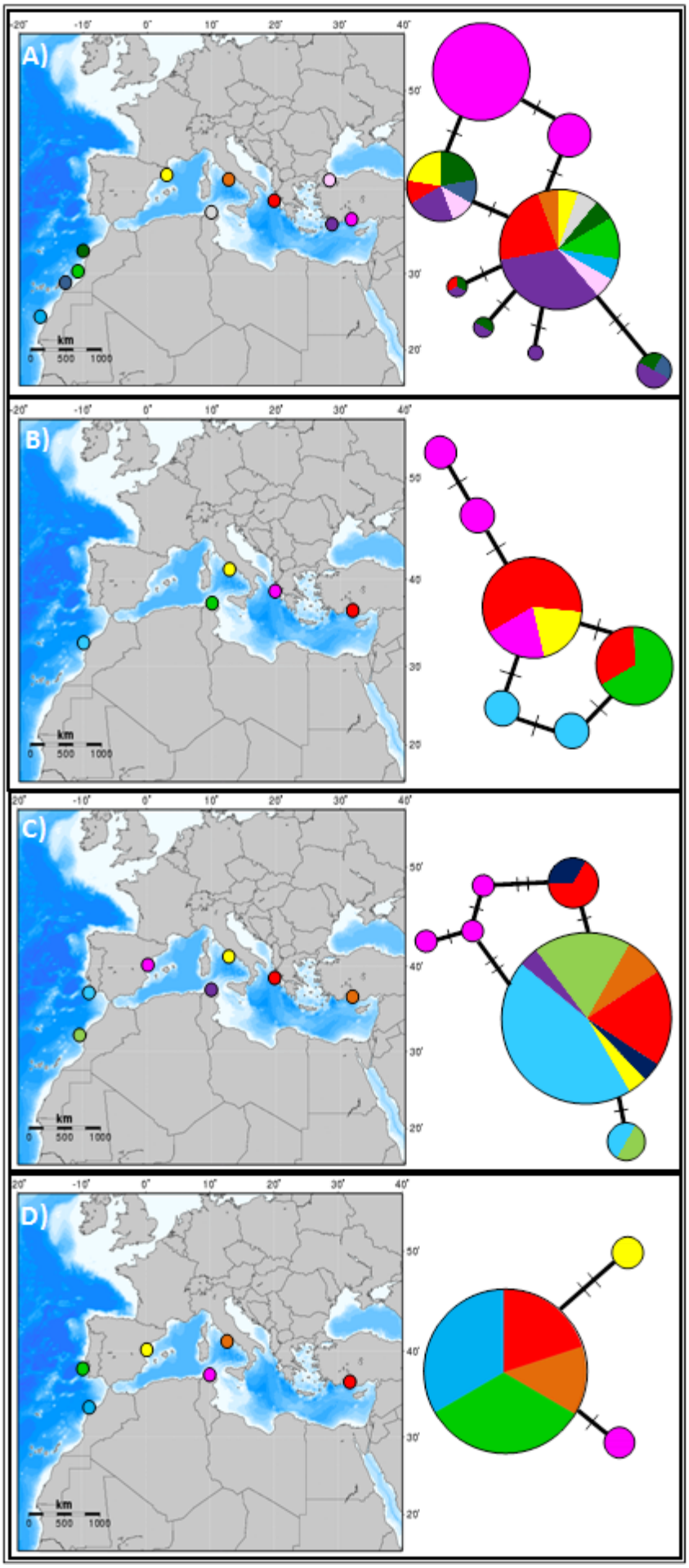


Figure 3.2.2: Median joining haplotype networks based on 509bp mtDNA COI for: A) *C. aurata;* B) *C. saliens;* C) *C. ramado; D*) *C. labrosus*. Node size is proportional to the haplotype frequency and node colour corresponds to sampling sites indicated on the adjacent map. Branch lengths are proportional to the number of sequence differences with each bar indicating a 1bp mutation.

3.2.3.2 Comparative phylogeography of Mugilidae across the Mediterranean-Atlantic divide

Sequences belonging to *C. aurata*, *C. saliens*, *C. ramado* and *C. labrosus* from locations within the Mediterranean basin and along the northeast Atlantic coasts of Morocco and Portugal were assessed for shared geographical patterns of genetic diversity and gene flow. Results for *M. cephalus* are presented separately below.

The RelTime analysis of relative sequence divergence (see Fig. 3.2.1) estimated intra-specific diversification of Mediterranean and northeast Atlantic taxa to have occurred since the mid Pleistocene, between ~1.17 MYA (*C. saliens*) and ~0.48 MYA (*C. ramado*). For *C. aurata* (Fig. 3.2.2A) 8 haplotypes were resolved across the 7 Mediterranean (including Sea of Marmara) and 4 Atlantic locations, whereas 6 haplotypes were resolved across the 4 Mediterranean and 1 Atlantic location for *C. saliens* (Fig. 3.2.2B) and the 5 Mediterranean and 2 Atlantic locations for *C. ramado* (Fig. 3.2.2C). Despite consisting of a similar number of samples from a comparable geographic area as the other species *C. labrosus* (Fig. 3.2.2D) exhibited a much lower overall haplotype diversity, with only 3 haplotypes resolved across the 4 Mediterranean and 2 Atlantic sampling locations.

Across samples from the Mediterranean and northeast Atlantic similar phylogeographic patterns were observed. For all four species MJ haplotype networks failed to identify any pattern of differentiation of haplotype composition or number between or within Atlantic and Mediterranean areas. A single common haplotype in the centre of the network was found in

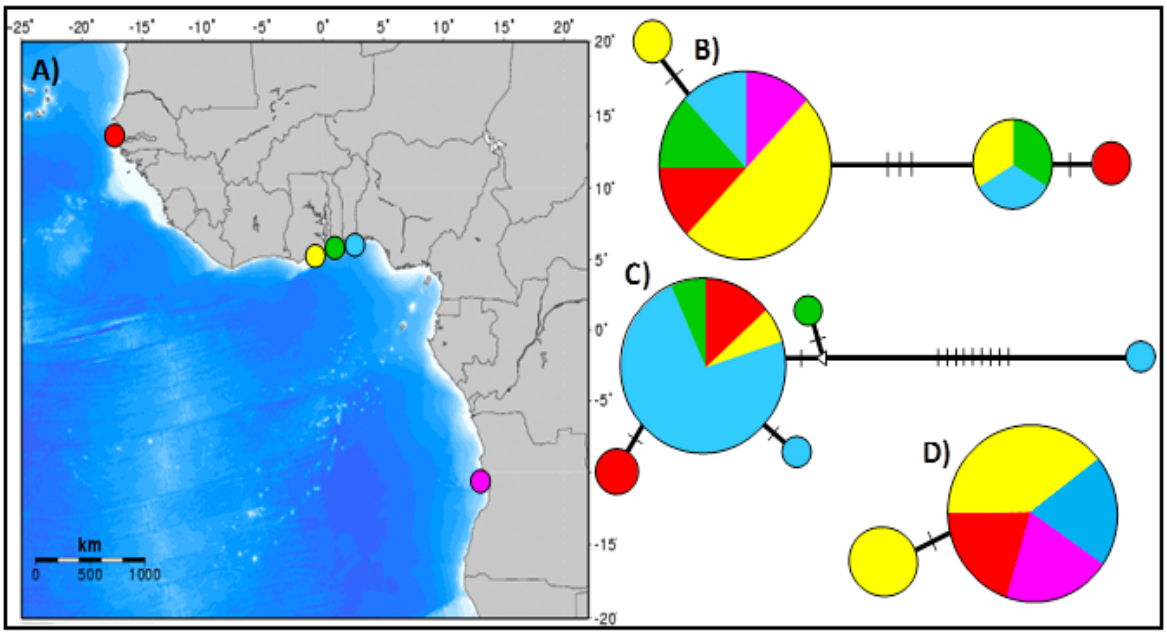


Figure 3.2.3: Median joining haplotype networks based on 509bp mtDNA COI for: B) *M. curema;* C) *N. falcipinnis;* D) *M. bananensis*. Node size is proportional to the haplotype frequency and node colour corresponds to sampling sites indicated on the adjacent map (A). Branch lengths are proportional to the number of sequence differences with each bar indicating a 1bp mutation.

the majority of individuals in both Atlantic and Mediterranean locations, with a second less common haplotype present in *C. aurata* and *C. saliens,* plus 2 (*C. labrosus*) to 6 (*C. aurata*) rare haplotypes, with no haplotype differing by more than 3 mutations from the central common haplotype. Private Atlantic haplotypes were observed only in *C. saliens* and *C. ramado*, and then only at very low frequencies.

3.2.3.3 Intra-specific comparative phylogeography of Mugilidae across the west African tropics

Across the tropical west coast of Africa, from Senegal to Angola, spatial patterns of gene flow and diversity were assessed for 3 species (*M. curema, M bananensis* and *N. falcipinnis*), although for N. *falcipinnis* no sequences were available to represent the southern tropics (Angola). *M. curema* and *N. falcipinnis* exhibited similar diversity, with 4 and 5 haplotypes resolved respectively, whereas *M. bananensis* displayed only 2 haplotypes, though this may be attributed in part to small sample size (n=6). All three species displayed a single common haplotype found in individuals from all locations, with other haplotypes at much lower frequencies. Both *M. curema* (Fig. 3.2.3B) and *N. falcipinnis* (Fig. 3.2.3C) exhibited some evidence of differentiation of 2 clades in the MJ topologies. For *M. curema* there appeared to be limited geographic basis to genetic structuring, with one clade containing sequences from all locations (Senegal, Ghana, Togo, Benin and Angola) and separated by 3 mutations from the second clade which also contained samples from all locations except Angola. The RelTime method placed the divergence of these 2 *M. curema* clades around 0.71 MYA (95%CI = 1.66-0). Stronger signals of genetic structuring were observed in the haplotype network of *N. falcipinnis*, where the first clade contained samples from all locations (Senegal, Ghana, Togo, Benin) and was highly divergent and separated by 9 mutations from the second clade, which contained a single sample from Ghana. RelTime estimates of node ages placed the divergence of these two *N. falcipinnis* clades around 4.32 MYA (95%CI=11.12-0).

3.2.3.4 Intra-specific comparative phylogeography of Mugilidae across the Benguela Upwelling System.

Patterns of genetic diversity across the BUS are only discussed for *C. tricuspidens*, however both broadly distributed species included in this investigation (*C. dumerili, M. cephalus*) incorporate samples from across the BUS and as such provide more depth to the analysis of this biogeographic barrier (see sections 3.2.3.4 and 3.2.3.5 below).

For *C. tricuspidens*, 3 haplotypes were resolved across the 1 Angolan and 1 South African location, with no sharing of haplotypes between the two locations (Fig. 3.2.4). The Angolan haplotype was distinct from the South African haplotypes by four mutations. Tamura-Nei genetic distances between haplotypes belonging to the South African and Angolan clades ranged

from 0.008-0.010. Estimated divergence time between these two clades placed the most recent common ancestor around 0.94 MYA (95% CI =2.29-0).

3.2.3.5 Phylogeographic patterns in *Chelon dumerili* across the tropical west coast of Africa and the Benguela Upwelling Current.

The broad distribution of *C. dumerili* across tropical west African and South African coasts allowed a more comprehensive phylogeographic investigation, employing sequences from northern Senegal along the west African coast to the far north east of South Africa.

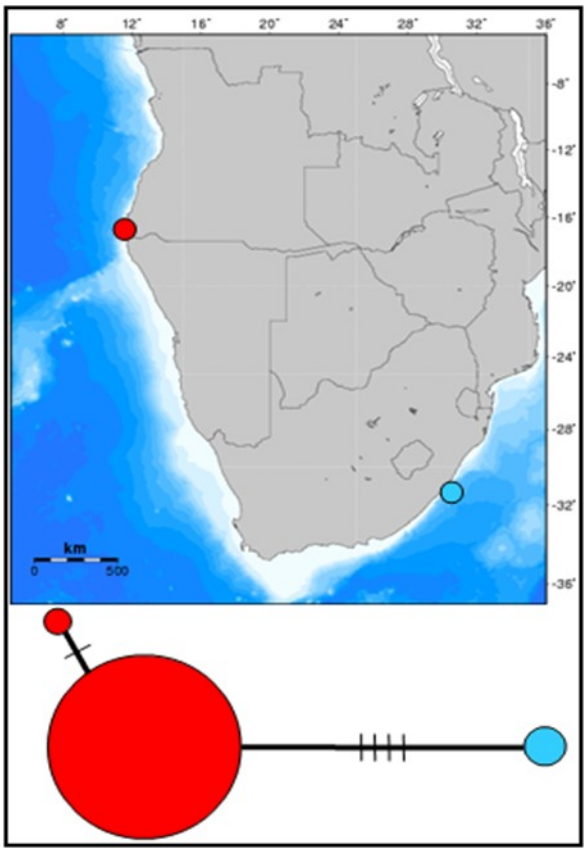


Figure 3.2.4: Median joining haplotype networks for *C. tricuspidens* based on 509bp mtDNA COI. Node size is proportional to the haplotype frequency and node colour corresponds to sampling sites indicated on the adjacent map. Branch lengths are proportional to the number of differences with each bar indicating a mutation.

For the COI dataset 7 haplotypes were resolved, which formed 3 highly divergent and reciprocally monophyletic clades (Fig. 3.2.5B & C– and see the wider ML tree in Fig.3.1.1 in Chapter 3.1). One clade consisted of all GENBANK sequences from Kwazulu-Natal (JF493799.1, JF493800.1, JF493801.1) and Great Fish River in South Africa (JQ060467.1, JQ060468.1), plus sequences generated in this study from Kwanza in Angola and the Kariega, Kowie and Mpambanyoni rivers in south eastern South Africa. This clade corresponds to the "Chelon spp. B" clade in Durand *et al.* (2012a). A second clade contained GENBANK sequences of individuals from the Gulf of Guinea (HM208826.1, HM208827.1), Senegal (JQ060463.1, JQ060464.1), Guinea-Bissau (JQ060466.1) and Togo (JQ060465.1), and corresponds with the "true *C. dumerili"* clade classified by Durand *et al.* (2012a). A third clade, unidentified by previous studies, consisted of sequences generated in this study from individuals from Kwanza in Angola. All three clades were highly diverged from one another, with the South African clade separated from the west African and Angolan clades by 30 and 35 mutations respectively, and the latter two clades divergent by 15 mutations. Accordingly Tamura-Nei genetic distances between haplotypes belonging to the South African clade and west African (T93=0.070-0.078) and Angolan clades (T93=0.081-0.089) were greater than those between the Angolan and west African clades (T93=0.030-0.034). Based on the average number of nucleotide differences between clades (South African-Angolan=37.27, South African-west African=32.31, Angolan-west

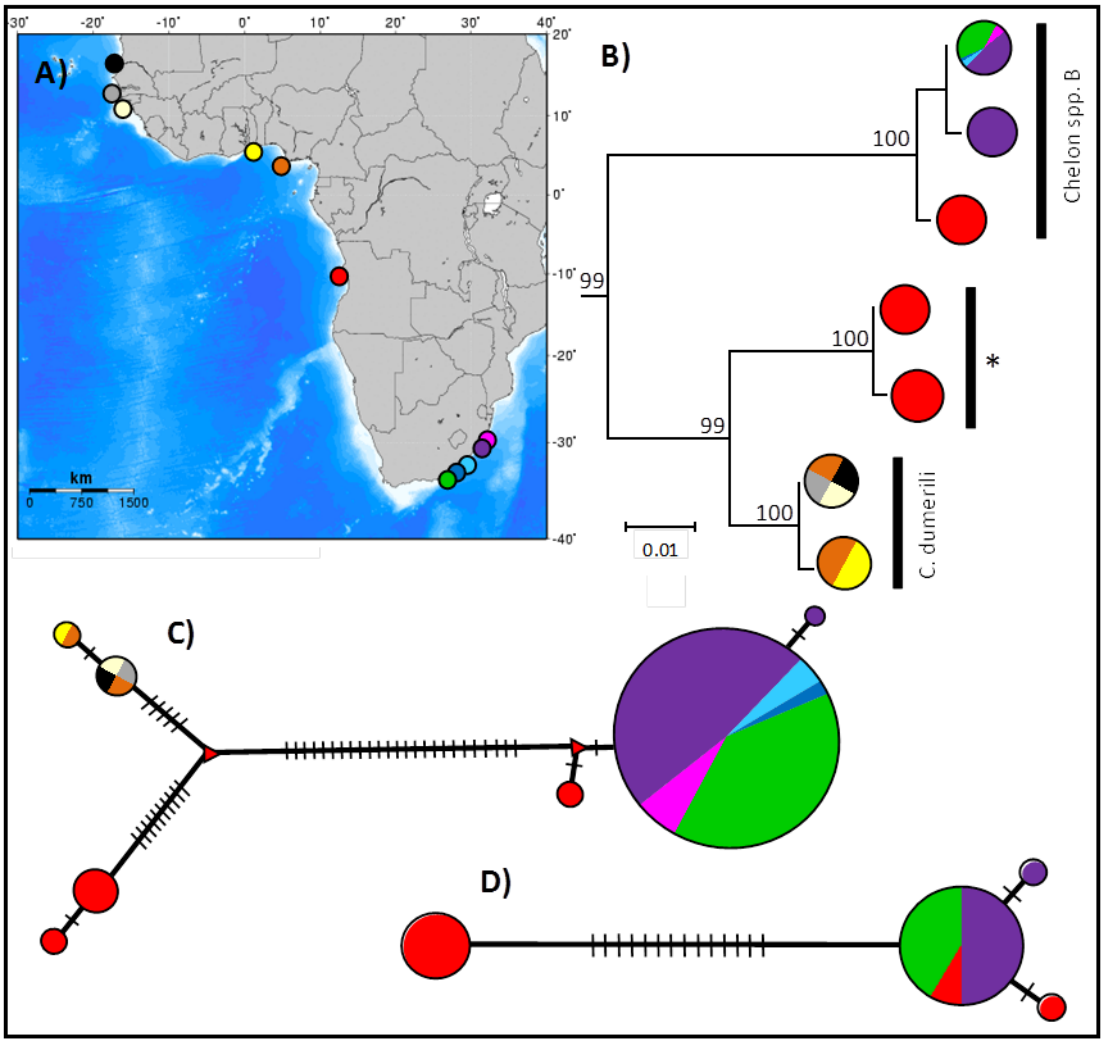


Figure 3.2.5: ML pylogeny (B) of C. dumerili based on 509bp of mtDNA COI, bootstrap values given above branches. Median joining haplotype networks for C. dumerili based on C) 509bp mtDNA COI and D)643bp of nuclear alpha-amylase intron 1. Node size is proportional to the haplotype frequency and node colour corresponds to sampling sites indicated on the map (A). Branch lengths are proportional to the number of differences with each bar indicating a mutation. * represents a novel lineage that has been resolved for the first time in the Angolan samples sequenced in this investigation.

African=15.62) percentage sequence divergence at COI was 3.07% between the West African and Angolan clades, with greater percentage sequence divergence observed between the South African and Angolan (7.32%) and west African (6.34%) clades. ML branching patterns (Fig. 3.2.5B) and RelTime node age estimates (Fig. 3.2.1) showed the South African clade (Chelon spp B) to split first from the most recent common ancestor around 7.31 MYA (95% CI=12.02-2.59), followed by a split between the west African (*C. dumerili*) and Angolan (*) clades around 3.4MYA (95%CI=6.88-0)

Due to the identification of a novel and highly diverged lineage (the Angolan clade) within *C. dumerili* a section of the nuclear alpha-amylase intron 1 was amplified for a subset of individuals, to confirm bi-parentally restricted gene flow. However as all sequences belonging to the west African clade were obtained from GENBANK, no representatives of this clade were included in the nuclear analysis and consequently only patterns of nuclear diversification be-

tween Angolan and South African individuals / clades can be examined. Sequences belonging to the South African clade contained 2 indels, one 22bp in length and the other 5bp, that were not observed within individuals belonging to the Angolan clade, so these were removed prior to all analyses. Within the nuclear alpha-amylase dataset no heterozygotes were detected and 4 haplotypes were resolved, which formed 2 highly divergent clades (Fig. 3.2.5D), corresponding to the South African and Angolan clades identified from the COI dataset. Again the South African clade consisted of samples from the eastern coast of South Africa, as well as 2 samples from Angola, whereas the Angolan clade contained only individuals collected in Angola. Both clades were highly diverged and separated from each other by 14 mutations.

3.2.3.6 Phylogeographic patterns in *Mugil cephalus* across the Mediterranean-Atlantic boundary, the tropical west coast of Africa and the Benguela upwelling system.

In total 14 haplotypes were resolved in the COI dataset, which fell into 3 highly divergent clades in the ML tree (Fig. 3.2.6). Clade 1 contained sequences obtained from fish inhabiting the Mediterranean basin and the northeast Atlantic coast of Morocco. Clade 2 comprised sequences from west African locations (Senegal, Togo, Benin, Angola) as well as a single sequence from Turkey, and clade 3 consisted solely of South African samples (Eastern Cape and Kwazulu-Natal). The western African and South African clades were separated from each other by 12 mutations, and both were separated from the Mediterranean clade by 8 or 17 mutations, dependant on treatment of Turkish haplotypes (discussed below).

The haplotype network (Fig. 3.2.6C) similarly defined 3 distinct haplotype clusters with clear separation of Mediterranean, Atlantic, tropical west African and South African groups. Both the South African and west African clusters were dominated by a single common haplotype which contained sequences from all (South African clade) or the vast majority (west African clade) of sampling locations. These common haplotypes were separated from less common or private haplotypes by 1 or 2 mutations for both clades. Conversely the Mediterranean-Atlantic clade displayed greater sequence diversity and can be further subdivided into 2 haplogroups/sub-clades. Haplogroup 1 consisted solely of GENBANK sequences from Turkish locations and was separated from each of the South African and west African clades by 8 mutations. Haplogroup 2 consisted of a single haplotype, which was more highly diverged from the South African and west African clades (separated by 17 mutations), and contained sequences from all Mediterranean and north East Atlantic locations (Turkey, Greece, France, Tunisia, Morocco). According to the ML plylogeny (Fig. 3.2.6B) all 3 clades (and 2 Mediterranean-Atlantic haplogroups) diverged from their MRCA nearly simultaneously, with the Mediterra-

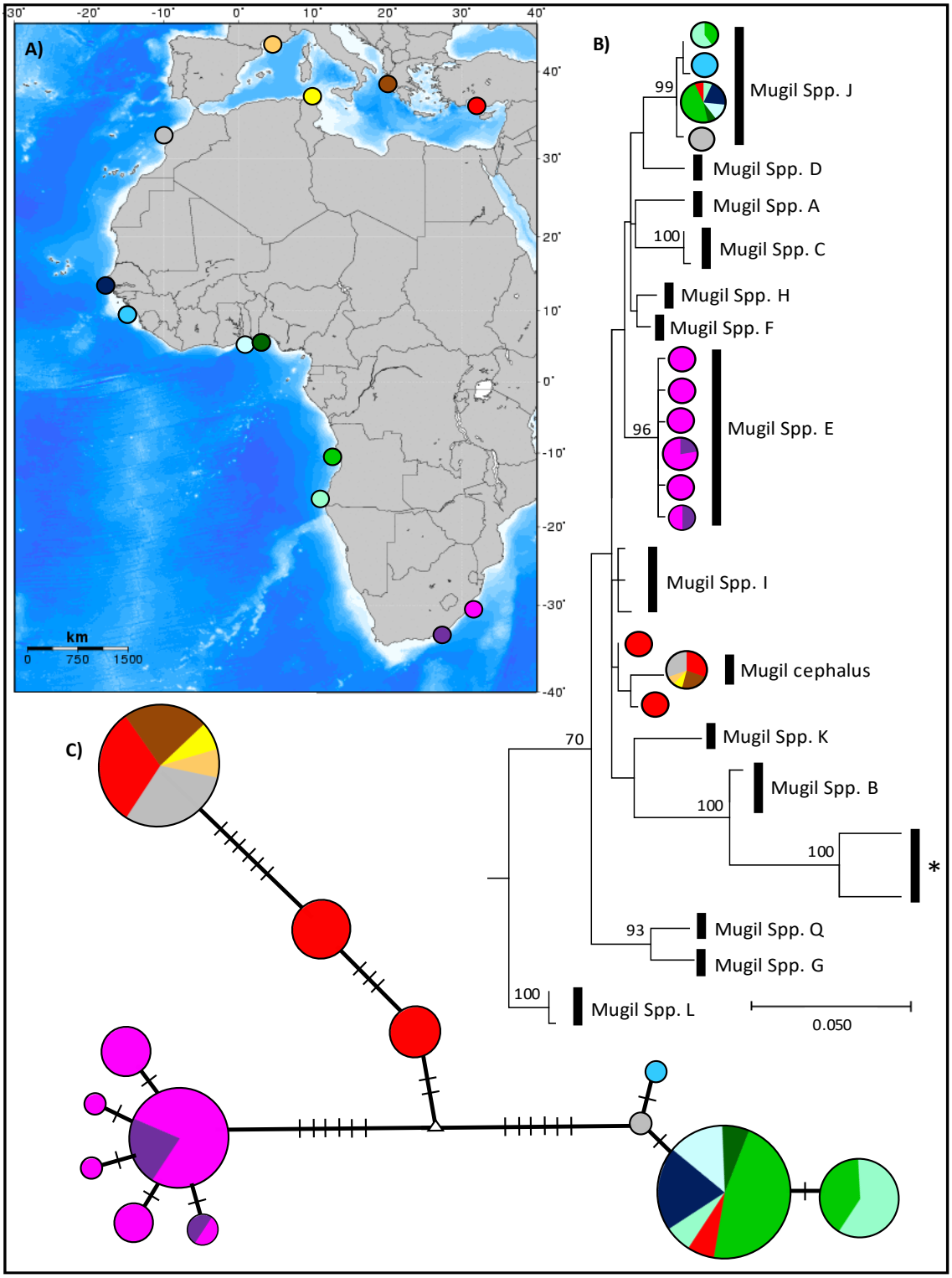


Figure 3.2.6: Maximum likelihood phylogeny (B) of all previously identified global *M. cephalus* lineages (see Durand 2017) based on 509bp of mtDNA COI, bootstrap values given above branches. * is the previously uncharacterised Egyptian clade (see section:3.1.3.1). Median joining haplotype network (C) for *M. cephalus* samples from the Mediterranean, west Africa and South Africa based on509bp mtDNA COI. Node size is proportional to the haplotype frequency and node colour corresponds to sampling sites indicated on the map (A). Branch lengths are proportional to the number of differences with each bar indicating a mutation.

nean clade (labelled *M. cephalus* in Fig. 3.2.6B) forming the basal position in the phylogeny. RelTime estimates of divergence dates within *M. cephalus*, placed this initial radiation around 4.94 MYA (95%CI = 8.48-1.39 MYA).

3.2.4 Discussion

The purpose of this study was to investigate shared patterns of genetic diversity, breakdown of genetic connectivity and speciation patterns in demersal fish (Mugilidae) across the Mediterranean, West African and South African coastlines in relation to recognised biogeographic boundary regions and potential barriers to dispersal. The Mugilidae provides an ideal family of fish in which to test these issues, due to the ecological similarity among the species as well as the number of species across the study area. Generally concordant patterns of genetic diversification across recognised biogeographic boundaries are shown in Mugilidae species; however phylogeographic patterns indicate subtle variations in evolutionary history and demography amongst species. Perhaps indicating some degree of stochasticity in the demographic history of species and their response to drivers of diversification, or a degree of uncharacterised ecological and/or life history variation.

3.2.4.1 Limited structure across the Mediterranean-Atlantic divide

Although the transition between the Atlantic Ocean and Mediterranean Sea has been highlighted as a significant phylogeographic boundary for numerous taxa (Patarnello *et al.* 2007), the present study found limited evidence of intra- or inter-specific genetic differentiation between Atlantic and Mediterranean populations of grey mullets. A similar lack of genetic structuring between the Atlantic and Mediterranean has been observed in many species, although it appears to be a pattern most prevalent in pelagic organisms with high dispersal potential (Ely *et al.* 2002; Pujolar *et al.* 2002; 2013; Daeman *et al.* 2001; Zardoya *et al.* 2004; Patarnello *et al.* 2007). Conversely, species that are reliant on estuarine or freshwater habitats for some of their life history, like grey mullets, often exhibit significantly more defined genetic structure between Atlantic and Mediterranean populations (Bernatchez 2001; Faria *et al.* 2012), so the grey mullets appear to be an exception to this pattern.

The shallow phylogenies and widespread distribution of common haplotypes, across the entire Mediterranean and the northeast Atlantic, for all four mullet species examined (excluding *M. cephalus*- see later) could be attributed to the ongoing high dispersal potential of these species, and/or recent expansion and colonisation of the region by these species. Although there has been no specific research on migratory behaviour and pelagic larval duration in *C. aurata*, *C. labrosus*, *C. saliens* and *C. ramado* the capacity to traverse great oceanic distances at both larval and adult life history stages has been observed in other species in this family.

The main migration of adult mullet is thought to occur between coastal areas and deeper water during a species' spawning period, and mark and recapture studies have recorded movement of adult mullet over hundreds of kilometres both within and outside of spawning periods (Idyll & Sutton 1952; Bernardon & Vall 2004; Bacheler *et al.* 2005). Mugilidae appear to also exhibit a relatively protracted pelagic larval duration of around 4 weeks (Li 1992), potentially facilitating long distance dispersal. Larval stage dispersal, however is governed by oceanographic currents (Chang *et al.* 2000), which across the transition area between the Atlantic and Mediterranean basins are characterised by an array of frontal zones (Almeria-Oran front - Tintore *et al.* 1988) and asymmetric flow of surface waters (Millot & Taupier-Letage 2005; Salat 1996). As such it is unlikely that the oceanography of the Atlantic-Mediterranean transition would promote active and symmetrical transport of, and consequent gene flow by, passively moving propagules between the Atlantic and Mediterranean basins, and so result in the lack of diversification observed here. An alternative explanation for the shallow phylogenies and lack of genetic structuring in mullets is that either Mediterranean or Atlantic populations suffered a localised extinction, followed by subsequent recolonization from the other basin, and/or an extreme bottleneck in both populations removing all but one or two haplotypes common across the species range (e.g. Bargelloni *et al.* 2005; Patarnello *et al.* 2007).

Based on geological records, the loss of ancestral Mediterranean clades due to unfavourable environmental conditions seems more likely than the loss of Atlantic populations. Pleistocene glacial maxima caused sea levels to drop by up to 170m (Nilson 1982), significantly reducing the depth and width of the Gibraltar Strait and possibly closing it completely. The consequent reduction in size and enclosed nature of the Mediterranean would therefore provide limited opportunity for species to track optimal environmental conditions, causing impoverishment of many Mediterranean taxa (Cunningham & Collins 1998). Furthermore the reliance of Mugilidae on estuarine and freshwater habitats could further limit their ability to track isotherms (Parmesan 2006). Optimal conditions for the re-invasion of the Mediterranean basin from the Atlantic could have been facilitated by interglacial sea level rises and the associated reconnection of these water masses. However without the sample sizes here to comprehensively test the historical demography of Atlantic and Mediterranean Mugilidae, inferences cannot be drawn on the direction of any re-colonization event. Nonetheless the recent geological timescale of this potential re-colonisation would provide little opportunity for the accumulation of genetic signatures of contemporary Mediterranean-Atlantic isolation and as such it is possible that for mullets, like in many other marine taxa, this biogeographic boundary does indeed represent a barrier to contemporary dispersal and gene-flow. The use of rapidly evolving and more polymorphic genetic markers (e.g. microsatellites) and larger sample sizes would be required in any future investigations to properly characterise patterns of contemporary gene

flow in Mugilidae.

Like other Mediterranean and north east Atlantic Mugilidae species, phylogenetic analysis of *M. cephalus* showed limited evidence of any disruption to gene flow in association with the Atlantic-Mediterranean divide. However phylogeographic structuring in *M. cephalus* within the Mediterranean basin was more complex, with 2 genetic groups resolved: a group represented by a single haplotype distributed throughout the Mediterranean, and a group found only in Turkish waters represented by two divergent haplotypes sitting phylogenetically between the first group (original *M. cephalus* according to Durand 2017) and the divergent W African and South African *M. cephalus* clades (Mugil species J and E, respectively, according to Durand 2017). This pattern contrasts with previous mtDNA genetic investigations of Mediterranean *M. cephalus* populations, where a single clade was reported across the entire Mediterranean, Black Sea and northeast Atlantic (Livi *et al.* 2011; Durand *et al.* 2013). Isolation and genetic differentiation of Aegean Sea populations from the rest of the Mediterranean has been observed for multiple species (Magoulas *et al.* 1996; Perez-Losada *et al.* 2007), with this genetic break attributed to lowered sea levels and population isolation during Pleistocene glacial maxima, with isolation maintained by contemporary oceanographic features of the region such as the quasi-circular anticyclonic front at the Aegean-eastern Mediterranean boundary (Malanotte-Rizzoli & Bergamasco 1989). Although limited mtDNA phylogeographic diversification was observed in a previous molecular investigation of *M. cephalus*, significant differentiation (i.e. isolation) at microsatellite loci was observed between Aegean, Black Sea and Mediterranean populations (Durand *et al.* 2013). However the distinct Turkish haplogroup resolved here consisted solely of samples from Yumurtalik (Keskin & Atar 2013), outside of the Aegean basin. Thus the observed patterns of differentiation could be the result of Aegean isolation, followed by a subsequent spread of Aegean haplotypes southward along the Turkish coast. An alternative explanation could be technical difficulties with the sequences from the Keskin & Atar (2013) study, as all previous studies found no such diversity, and as such a re-examination of *M. cephalus* from this area would be desirable.

3.2.4.2 Comparative phylogeography of Mugilidae across the West African tropics

Patterns of phylogeographic diversification in Mugilidae were not congruent among species with respect to the biogeographic provinces of west Africa specifically. The different patterns resolved were: (i) no divergence across the west African coast from Angola to Senegal in *M. bananensis* and *M. cephalus.* (ii) divergence of Angolan samples from all locations further north in *C. dumerili* (discussed further in section 4.2.4). (iii) more complex patterns of non-geographically correlated intra-specific genetic structuring resolved in *M. curema* and *N. falcipinnis*.

The lack of genetic structure reported across tropical West Africa in *M. cephalus* and *M. bananensis* most likely reflects persistence in a common historical refuge with no restriction to gene flow post-colonisation from this refugial area. This explanation is supported by the west-African refugia hypothesis (Maley 1991; Lévêque, 1997), which proposes that fragmented regions of tropical conditions (warm and wet) represented refugia for many species in the overall aridity of western Africa during glacial maxima, and that the presences of these refugia shaped the phylogeographic patterns of estuarine-associated aquatic fauna. Major core areas of tropical refugia have been identified in Angola, Cameroon/Gabon, western Ghana/Ivory coast and Sierra Leone/Gabon (Maley 1991). However to determine the most likely refugia of *M. cephalus* and *M. bananensis* an extensive analyses of genetic diversity and demographic history would be required, and as such would necessitate a more comprehensive sample set than was implemented here.

In contrast to the shallow phylogenies of west African *M. cephalus* and *M. bananensis* populations, in *M. curema* two clades were resolved across tropical West Africa with limited obvious geographical association. Similar patterns of genetic structuring, with no geographical clustering of haplotypes, were resolved for west African bluefish (*Pomatomous saltatrix*; Reid *et al.* 2016), a pattern attributed to either stochastic coalescence or historical habitat fragmentation followed by secondary contact of isolated populations. In the case of *M. curema*, reported levels of divergence between clades and the timescale estimated (0-1.66 MYA) for this divergence event align more readily with the latter explanation of historical fragmentation and secondary contact. Similar to other Mugilidae species investigated here, initial isolation was likely generated by Pleistocene environmental vicariance and its impact on estuarine and aquatic habitats (as reported for other tropical west African estuarine species: Falk *et al.* 2003; Durand *et al.* 2005b; 2013).

Finally, although analysis of *N. falcipinnis* lacked an Angolan sample, this species also appeared to follow a similar pattern of low diversity and region-wide genetic homogeneity as seen in *M. bananensis*, *M. cephalus* and *M. curema*. There was, however, identification of a highly divergent lineage within Ghanaian samples of *N. falcipinnis*. A similar pattern of diversification of Ghanaian samples from other closely situated west African samples was reported in *Sardinella aurita* (Chikhi *et al.* 1998), although in this case observed patterns were attributed to sampling error and/or variances in reproductive success. Genetic distances between the divergent *N. falcipinnis* Ghanaian clade and other West African haplotypes exceeded accepted barcoding gaps (within species T93 distance=0.002-0.006; between-clade T93 distance=0.021-0.023), and were greater than previously described thresholds for delineation of cryptic Mugilidae species (Durand & Borsa 2015). As such, and pending nuclear genetic investigation, the presence of

previously uncharacterised cryptic (possibly species-level) diversity within *N. falcipinnis* should be noted.

Despite the small sample sizes employed here and the reliance on a single locus to resolve phylogenetic relationships, these preliminary results constitute an important contribution to the wider understanding of the evolutionary history of tropical west African species. The region extending from Senegal to Angola remains remarkably under-investigated, particularly in relation to finer scale evolutionary history and population structure of the species that inhabit it. As many west African fisheries are approaching overexploitation (Gascuel *et al.* 2007; Atkinson *et al.* 2011), gaining an understanding of the spatial distribution of genetic diversity is fundamental for the establishment of sustainable management and maintenance of healthy stocks.

3.2.4.3 Comparative phylogeography of Mugilidae across the Benguela Upwelling System

All three Mugilidae species (*C. dumerili*, *C. tricuspidens* and *M.cephalus*) described as spanning the BUS revealed highly diverged and reciprocally monophyletic lineages between Angola and South Africa. These phylogeographic breaks in all three species correspond with the area of the Namibian coast which is dominated by the large perennial upwelling cell (the Luderitz upwelling) generated by the Benguela Current.

The contemporary oceanographic features of the BUS are thought to have been established during the Plio-Pleistocene transition (~2 MYA - Marlow *et al.* 2000; Krammer *et al.* 2006), although the current itself became fully established much earlier around 6 MYA (Shannon 1985). Timescales of divergence between populations isolated to either side of the upwelling cell have been associated either with the original formation of the upwelling system (Bowen *et al.* 2001; Teske *et al.* 2007; Floeter *et al.* 2008; Henriques *et al.* 2014; Henriques 2012; Gwilliam 2017) or subsequent environmental perturbations during Pleistocene glacial and inter-glacial cycles (Henriques *et al.* 2012; Reid *et al.* 2016). Divergence date estimates placed the diversification of the South African and Angolan clades of *C. dumerili* far earlier (~7.31 MYA) than the equivalent clades in *C. tricuspidens* (~0.94 MYA). The estimates for *C. dumerili* agree with divergence dates previously reported between South African and west African taxa (Bowen *et al.* 2001; Teske *et al.* 2007; Floeter e*t al.* 2008; Henriques *et al.* 2014; Henriques 2012; Gwilliam 2017), including *C. dumerili* clades (Durand & Whitfield 2015). The divergence between Angolan and South African *C. tricuspidens* clades falls within the Pleistocene epoch, and as such was most likely promoted by environmental changes during glacial and interglacial cycles, as has been attributed to intra-specific genetic structuring across the BUS in previous investigations (Henriques *et al.* 2012; Reid *et al.* 2016). For *M. cephalus* however a rapid

and near simultaneous global radiation of all lineages within this species complex, which appears to have occurred near simultaneously during the Pliocene epoch, obscures fine scale assessment of the role of the BUS in shaping the phylogeographic partitioning of this species.

Furthermore, all the mullet species investigated here are catadromous, spawning at sea but relying on estuarine and freshwater habitats for much of their life cycle (Blaber 1977). Catadromous life history characteristics have important influences on the patterns of spatial genetic heterogeneity, however the extent of which is thought to be tightly linked to spawning behaviours (e.g. distance from estuary mouth) and larval dispersal (Avise et al. 1986; Sang et al 1994; Chenoweth et al. 1998; Jerry & Baverstock 1998). Although very little is known about the spawning behaviour of many Mugilidae species, available evidence suggests an extensive dispersal potential in adult mullets (Idyll & Sutton 1952; Bernardon & Vall 2004; Bacheler *et al.* 2005), and as such high gene flow, similar to other migratory catadromous species (e.g. *Anguilla japonica*—Han et al. 2010; *Angulla anguilla*– Avise et al. 1986; Dannewitz et al. 2005) would be expected. Consequently it is likely that the resolved break down in gene flow between Angolan and South African mullets reported here can be attributed to geographic features (i.e. the BUS) rather than life history per se.

A further confounding factor related to the reliance of Mugilidae on estuarine habitat is disentangling the relative influence of the BUS verses available estuarine and freshwater habitats on the resolved genetic partitioning. The overall aridity of the Namibian coast, where the upwelling cells of the BUS are most intense, has resulted in a distinct lack of suitable freshwater and estuarine habitats, which are fundamental resources for juveniles and adults of the majority of African Mugilidae species (Blaber 1977). As a result of this lack of suitable habitat, great dispersal distances would be required (either actively as adults or passively as larvae) to maintain high gene flow between South African and Angolan population, with no stepping stone habitats available. It is therefore likely that in combination with the oceanographic features of the BUS, contemporary isolation of Mugilidae populations to the north and south of Namibia is being maintained by this lack of suitable estuarine habitats in Namibia. This is supported by previous population genetic investigations of estuarine associated species, where gene flow between estuaries followed a one dimensional stepping stone model (Kimura & Weiss 1964) and conformed to a scenario of IBD (Chenoweth et al. 1998; Jerry & Baverstock 1998).

3.2.4.4 Identification of a novel mtDNA lineage and patterns of phylogeographic diversification across west Africa and South Africa in *C. dumerili*

A key finding of this investigation was the identification of a previously undescribed lineage within the *C. dumerili* species compex. Previous phylogenetic investigations of the *C. dumerili* species complex (Durand *et al.* 2012a; Durand & Borsa 2015; Durand *et al.* 2017; Durand & Whitfield 2015; Xia *et al.* 2016) identified only 2 lineages: one occurring along the eastern coast of South Africa (Chelon species B) and the other occurring along the west African coast from the Gulf of Guinea north to Senegal (*C. dumerili*). The present study identified the two described lineages plus a third highly divergent and reciprocally monophyletic lineage found solely in Angolan samples. Genetic distances between *C. dumerili* lineages were greatest between the South African lineage and the Angolan (TN93 = 0.081-0.089, sequence divergence = 7.32%) and west African (TN93 = 0.070-0.078, sequence divergence = 6.34%) lineages, whereas the Angolan and west African lineages were more genetically similar (TN93 = 0.030-0.034, sequence divergence = 3.07%), with levels of intra-specific genetic distance not exceeding 0.008 for any species. Following Durand & Borsa (2015) the threshold for delineating species within the Chelon genus based on barcoding gaps is 1%, furthermore the level of divergence observed here amongst all three *C. dumerili* lineages exceeds those between other recognised Mugilidae species and cryptic species. As such (and pending nuclear genetic conformation) the Angolan, west African and South African lineages resolved here should be considered distinct sister species within the *C. dumerili* species complex.

ML phylogenetic analysis and corresponding divergence date estimates (Fig.3.2.4 & 3.2.6) suggest that the South African and Angolan/west African lineages diverged first around 7.31 MYA (in accordance with the 7.7 MYA date given to this node in Durand & Whitfield 2015) and so coinciding with the formation and establishment of the BUS around 6MYA (Krammer et al. 2006). Subsequently the west African and Angolan *C. dumerili* lineages diverged, with this divergence estimated as occurring around 3.4 MYA. As such the isolation of these lineages in Angola and western Africa is most likely attributed to the onset of pronounced glacial/ interglacial cycles during the Pliocene-Pleistocene transition. The isolation of populations at the far southern limits of the west African tropics (Angola) from those occurring further north has been observed in other estuarine-associated species (Durand *et al.* 2005b; 2013; Falk *et al.* 2013) and can be attributed to vicariant refugia during glacial maxima. Globally the phylogeographic structure and demographic history of species that are reliant on estuarine habitats has been shown to be heavily influenced by Pleistocene glaciations (Beheregaray and Sunnucks 2001; Beheregaray *et al.* 2001; Burridge *et al.* 2004; Chenoweth *et al.* 1998a,b; Fauvelot *et al.* 2003). In tropical western Africa, it is likely that glacial and interglacial cycles saw cyclic changes between periods of extreme aridity during glacial maxima and more wet tropical con-

ditions during interglacials (Lévêque 1997). Geographic areas exhibiting high present day endemism have been identified as potential refuge zones for estuarine taxa during glacial cycles (Maley 1991; Lévêque 1997), with one such refuge zone identified in Angola. Consequently the isolation of Angolan populations of *C. dumerili* is likely a relic of a tropical Angolan glacial refuge, which became isolated from populations further north and subsequently evolved allopatrically.

The final novel feature within the *C. dumerili* phylogeny (Fig.3.2.4 B&C) was the occurrence of the South African lineage among Angolan samples, a result corroborated by the nuclear genetic analysis (Fig. 3.2.4D). This result indicates some degree of asymmetrical dispersal (from South Africa to Angola) and secondary contact across the BUS. It is therefore likely that the BUS does not represent an entirely impermeable barrier to dispersal for *C. dumerili*. The strength of the upwelling cells within the BUS are seasonally influenced, with anomalous weakening of upwelling cells and intrusion of tropical waters further south occurring several times in the last century (Taunton-Clark & Shannon 1988; Stander & De Decker 1969; Gammelsrod *et al.* 1998). Accordingly, bi-directional gene flow between South Africa and Angola has been recorded in *Argyrosomus inodorus* (Henriques *et al.* 2015). However it would appear that for *C. dumerili,* as suggested in *Pomatomous salatrix* (Reid *et al.* 2016), dispersal occurs asymmetrically from South Africa to Angola, likely facilitated by anomalous warming events in combination with the general northward flow of the Benguela Current.

3.2.4.5 Phylogeography and patterns of speciation among Mugilidae species across the Mediterranean-NE Atlantic, tropical west African and southern African biogeographic regions

MtDNA phylogeography among Chelon species distributed across eastern Atlantic coasts support a process of vicariance-driven speciation consistent with proposed biogeographic boundary regions within the area. The deepest inter-specific divergences occurred between all northeast Atlantic / Mediterranean species (*C. aurata, C. ramado, A. labrosus* and *C. saliens*) and those occurring further south in the east Atlantic and the southwest Indian Ocean. Divergence dates, based on fossil calibrations of the ML phylogeny, place the separation of northeast Atlantic taxa in the early- to mid-Miocene (16-9 – 12.7 MYA). Similar Miocene epoch speciation of northeast Atlantic from tropical eastern Atlantic taxa has been reported in Sparid fish (Santini *et al.* 2014), and is consistent with vicariant speciation during the Mid-Miocene-Climatic transition (MMCT), a period in which the intensity of the CCUS was strengthened (Emery & Uchupi 1984).

Subsequent to the mid-Miocene split of north east Atlantic from west African and southern African species a divergence event forming two sister lineages occurred around 11.8 MYA. The

first of which's contemporary distribution is restricted to the south west Indian Ocean (Chelon spp A) and the cool temperate and temperate coast of South Africa (*C. richardsonii*), with the second lineage comprised of two species found in the tropics (*C. bandialensis*) and around the BUS (*C. tricuspidens*). Considering the clear intra-specific partitioning of *C. tricuspidens* haplotypes to the north and south of the BUS, it is likely that *C. tricuspidens* was restricted either to the coast of South Africa or, more concordant with the results here, restricted to the tropical eastern Atlantic. In this case, divergence between tropical west African taxa and southern African / Indian Ocean species can be readily aligned with the formation and strengthening of the southeast Atlantic upwelling cell (the early stages of the BUS), which occurred around 12 MYA (Siesser 1980; Diester-Haass *et al.* 1990; Krammer *et al*. 2006).

An alternative explanation for the divergence between tropical and South African/Indian Ocean Chelon species could be the steepening of thermal gradients between the tropics and the poles due to expansion of polar ice-sheets during the mid-Miocene (Norris 1998), which served to isolate temperate species in the northern (C*. bandialensis, C. tricuspidens*) and southern (Chelon spp. A, *C. richardsonii*) hemispheres. In this scenario post-divergence founder events from the northern hemisphere populations to southern Africa could explain the contemporary distribution of *C. tricuspidens* (e.g. Ludt *et al.* 2018). Trans-equatorial founder events would equally explain the subsequent divergence of *C. tricuspidens* (southern Africa) from the northern tropical east Atlantic *C. bandialensis* (e.g. Grant & Leslie 2001; D'Amato *et al.* 2008). Alternatively fossil records and isotopic data suggest that during the Miocene-Pliocene transition the tropical coast of west Africa was substantially warmer than present (Yemane *et al.* 1985; Cerling *et al.* 1997). Therefore a late Miocene to early Pliocene divergence between *C. bandialensis* in the northern tropical east Atlantic and *C. tricuspidens* in the southern tropical east Atlantic could also align with an isolation of these species at either side of the tropics.

The reliability and accuracy of divergence date estimates based on molecular clocks have been called into question in numerous investigations (e.g. Ho *et al.* 2005; 2011; Welch & Bromham 2005; Ho & Phillips 2009; Warnock *et al.* 2011; 2017; Duchêne *et al*. 2014; Heath et *al.* 2014; Donoghue & Yang 2016), and as such placing considerable emphasis on interpretations based on these estimates should be cautioned. Regardless of the absolute timing of these divergence events, the order of diversification and the relative timings of these events, in combination with the contemporary distribution of these species suggest a prominent role for vicariance-driven speciation in Mugilidae across the eastern Atlantic, that can be associated with the major biogeographic boundaries of this region (i.e. the CCUS, tropical equatorial west Africa and the BUS). Furthermore, the identification of persistent isolation of northeast

Atlantic species from those further south emphasise the importance of the CCUS in shaping the evolutionary history of the Mugilidae.

3.2.4.6 Conclusion

Prior to the present investigation patterns and levels of within-species divergence and diversity had been investigated comprehensively in only a handful of Mugilidae species, most prominently *M. cephalus* (Crosetti *et al.* 1993, 1994; Heras *et al.* 2009; Jamandre *et al.* 2009; Ke *et al.* 2009; Liu *et al.* 2009; Livi *et al.* 2011; Rocha-Olivares *et al.* 2000; Rossi *et al.* 1998b; Shen *et al.* 2011; Durand *et al*. 2013) and *M. curema* (Fraga *et al.* 2007; Heras *et al.* 2006, 2009). As such this study represents the first assessment of intra-specific patterns of evolutionary diversification for a wide range of Mugilidae species, and provides a basis for future phylogeographic and population genetic investigations of grey mullets. Through assessment of the spatial distribution of mtDNA genetic diversity, this study found grey mullets to exhibit a wide range of genetic structures and inferred evolutionary responses to recognised biogeographic boundaries across the Mediterranean, west Africa and South Africa. Structuring included concordant patterns across multiple species of genetic homogeneity across the Atlantic-Mediterranean transition, persistent and prolonged species level diversification across the CCUS and parallel patterns of intra- and inter- specific phylogenetic diversification across the BUS, as well as more complex species-specific patterns of vicariance-driven isolations, localised extinctions and secondary contact along tropical west African coasts and estuaries.

Although for many species included in these analyses sample size inhibited the ability to more rigorously characterise and test the drivers of observed phylogeographic patterns, commonalities in results suggest to some degree the presence of shared drivers of diversification across species and geographic regions. The only shared phylogeographic partition observed across all intra-specific comparisons of Mugilidae species was in association with the BUS, highlighting the capacity of large areas of cold upwelled water to influence the distributional patterns and consequent evolutionary diversification of Mugilidae species over long evolutionary timescales. An alternative way in which to assess the impact of these proposed biogeographic boundaries on the Mugilidae is to examine the proportion of trans-boundary species within and across genera (e.g. Floeter *et al.* 2007). Accordingly, the CCUS coincides with the far northern/southern distributional limits of the majority of west African and NE Atlantic Mugilidae species/lineages (respectively), and also with a phylogeographic/phylogenetic break between Mediterranean-NE Atlantic and west African clades/species within the Chelon genus and *M. cephalus* (the only Mugilidae species whose distribution extends across the CCUS). Within the Chelon genus ML topologies resolved a deep and persistent isolation of northeast Atlantic species from those distributed further south (i.e tropical west African and southern Africa), which

further emphasises the significance of the CCUS in shaping inter-specific genetic structuring of the Mugilidae.

It would appear that a large proportion of intra-specific phylogeographic diversification in Mugilidae may be linked to climate-associated drivers of population demography and connectivity, either historically during Pleistocene glacial and interglacial cycling or because of contemporary water temperature regimes, with most intra-specific diversification estimated to have occurred within the last million years. The same linkages may be evident also in shared patterns of phylogeographic distribution amongst species, as concordance in limitations to, and patterns of, distribution amongst species can be indicative of inherent oceanographic and geological forces that have disrupted gene flow and promoted speciation. Mediterranean, west African and South African Mugilidae species generally conform to a northern temperate (Mediterranean and NE Atlantic), tropical west African (Senegal to Angola), southern temperate (Namibia and South Africa) or tropical Indian Ocean / Indo-Pacific distribution. The apparent affiliation between the distribution of grey mullet species and the temperature of their environment can be readily attributed to life history and ecological characteristics inherent to this family, such as reliance on estuarine habitat and the influence of water temperature on spawning behaviour and timing (Bruslé 1981; Kelley *et al.* 1991; Kuo 1995). Ultimately the apparent vulnerability of Mugilidae to population fragmentation, localised extinctions and distributional shifts in response to changing environmental temperatures highlights the risk of future climate change to Mugilidae populations, and as such should be considered in future management and conservation objectives.

CHAPTER 4

Evolutionary history of the *Trachurus trachurus* species complex across the eastern Atlantic

CHAPTER 4

Genetic analysis provides insights into species distribution and population structure in east Atlantic horse mackerel (*Trachurus trachurus* & *T. capensis*).

4.1 Introduction

The horse/jack mackerels are a widely distributed genera of Carangid fish found throughout temperate, tropical and subtropical oceanic and coastal waters (Fricke *et al.,* 2019). Across the temperate waters of the north east (NE) Atlantic, Mediterranean and southern Africa two sister species, *Trachurus trachurus* and *Trachurus capensis*, are described based on morphological (Nekrasov, 1976; but see: Froese & Pauly, 2001) and genetic (Cardenas *et al.,* 2005; Karaiskou *et al.,* 2004) characteristics. *T. capensis* (Catelnau, 1861) is associated with the Benguela Upwelling System (BUS) and is most often described as extending from southern Angola, throughout Namibia and into the Eastern Cape of South Africa (Axelsen *et al*., 2004; Barange *et al*., 1998). *T. trachurus* is widely described throughout the NE Atlantic as far north as Norway and Iceland (Knijn *et al*., 1993) as well as the Mediterranean basins (Fischer, 1973). Both species are of great commercial importance and are the most valuable horse mackerel species in these regions (Murta, 2000) as exemplified by the Namibian *T. capensis* fishery which reports the highest biomass and catch of all fish species landed within the region and is the second highest economic contributor to Namibian fisheries (Kirchner *et al.,* 2010).

T. trachurus has demonstrated severe reductions in total annual catch with the stocks in the Mediterranean, Alboran and Black seas considered fully exploited, while the NE Atlantic stock is deemed to be overexploited (ICES, 2003, 2005). This has led to *T. trachurus* being listed as 'vulnerable' on the IUCN's red list of threatened species (Smith-Vaniz *et al.,* 2015) and added further impetus to align management and biological units for both species to try and ensure fishery sustainability (reviewed in: Abaunza *et al*., 2008a) and maintenance of biodiversity. Genetic studies of *T. trachurus* employing mitochondrial DNA (Comesaña *et al*., 2008; Karaiskou *et al*., 2003; 2004), allozyme (Cimmaruta *et al*., 2008) and microsatellite (Kasapidis & Magoulas, 2008; Sala-Bozano *et al*., 2015) markers have typically reported a general lack of genetic differentiation across NE Atlantic and Mediterranean samples. In the absence of robust evidence of population structure the species is managed using regionally designated stocks (ICES, 2003, 2005). Similarly, population genetic studies of *T. capensis* have reported no evidence of population structure (Karaiskou *et al.,* 2004; Naish, 1990) with the species currently managed as arbitrarily delineated northern and southern BUS stocks (Hecht, 1990; Naish *et al*., 1991).

Studies to date have focused on either the NE Atlantic/Mediterranean or the south west African coast meaning that *T. trachurus* and *T. capensis* diversity within the Angolan/North African region is poorly resolved. Such a geographical sampling gap also adds ambiguity as to the southern and northern limits of *T. trachurus* and *T. capensis*, respectively. Although the majority of studies cite the distribution of *T. trachurus* as extending only as far south as Senegal and the Cape Verde islands (e.g. Abaunza *et al*., 2008a; Cimmaruta *et al*., 2008; Comesaña *et al*., 2008; Sala-Bozano *et al*., 2015), others recognise a far wider range, encompassing the west African tropics and ranging as far south as South Africa and southern Mozambique (Boateng, 2019; Koranteng, 2002; Smith-Vaniz *et al*., 2015). Considering the extensive fishing pressure upon *T. capensis* in Namibia, the potential for overlap, and thus a cryptic mixed stock with *T. trachurus* must be assessed. Additionally, the available genetic data does not sufficiently test the validity of the currently recognised north and south BUS stocks for *T. capensis*.

To address these fundamental knowledge gaps for both species this study focused on collecting samples of putative *T. trachurus*/*T. capensis* from the understudied west African coast alongside better studied, and thus reference, NE Atlantic/Mediterranean (*T. trachurus*) and South African (*T. capensis*) waters. Genetic analysis of sequence variation at two mtDNA regions and 10 microsatellite loci were used to delineate species and assess population structure within both species. The results provide new insight into the distribution of both species and their eco-evolutionary dynamics while also providing information as to the relationship between population genetic structure and current management units.

4.2 Materials and Methods

4.2.1 Sample acquisition and DNA extraction

This study complied with all ethical requirements of the Journal of Fish Biology and local authorities. No fish were killed or interfered with in any way as part of this study as all samples were in the form of ethanol preserved fin clips of adults acquired solely from artisanal fish markets to ensure local providence of individuals. Samples were acquired from locations (Table 1 and Figure 1) across the western Mediterranean basin (Morocco: Al Hoceima), the Northeast Atlantic (Portugal: Olhao, Morocco: Essaouria and Tiznit), the tropical/subtropical Atlantic coasts of West Africa (Ghana: Tema, Angola: Namibe), and the cool temperate west coast of South Africa (Saldanha) between July and September 2017.Total DNA was extracted using a standard CTAB-chloroform/isoamyl alcohol method(Winnepenninckx *et al.,* 1993).

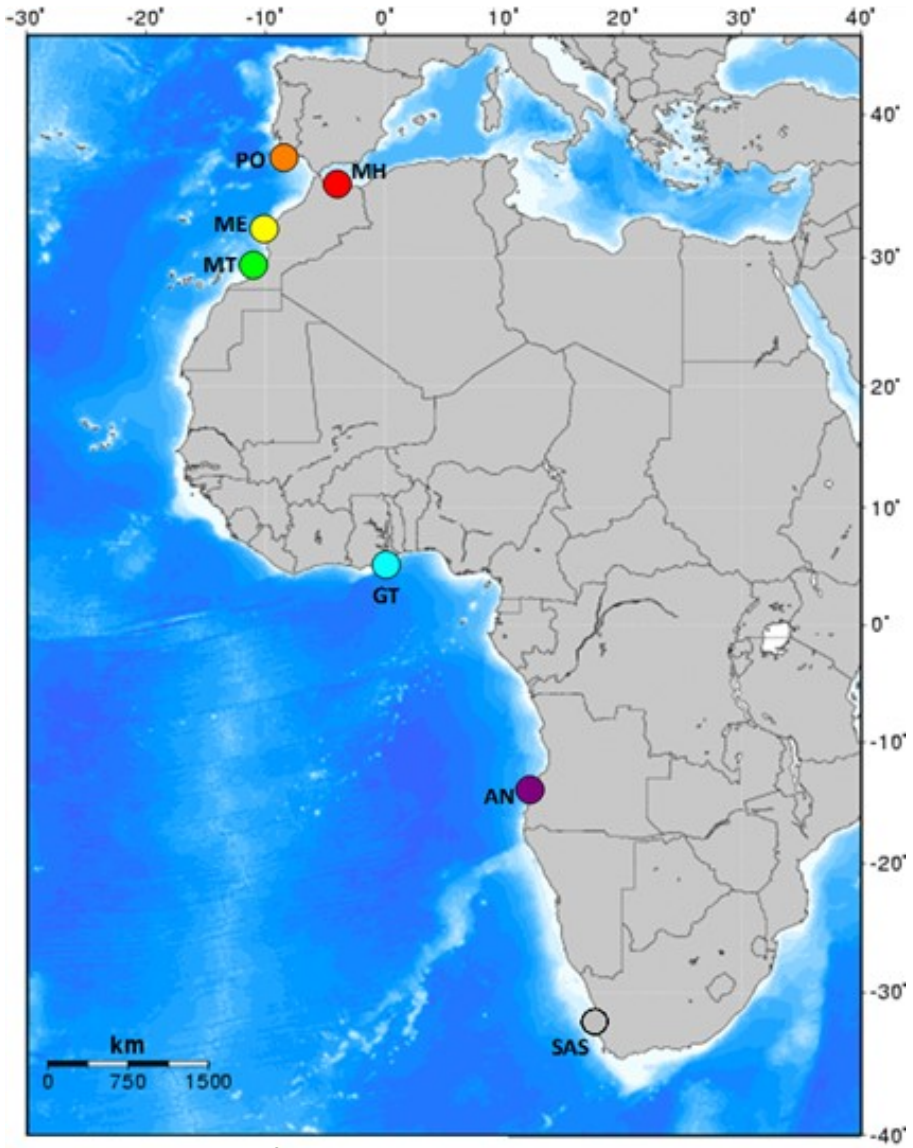


Figure 4.1: Sampling locations of horse mackerel across the Mediterranean and east Atlantic. MH= Morocco, Al Hoceima; PO= Portugal, Olhao; ME=Morocco, Essaouria; MT= Morrocco, Tiznit; GT= Ghana, Tema; SAS= South Africa, Saldanha.

4.2.2 MtDNA amplification & analysis

Two mtDNA regions were sequenced. A fragment of the cytochrome oxidase subunit 1 (COI) was PCR amplified using the Fish F1 and Fish R1 universal primers (Ward *et al.,* 2005) and a fragment of the Control Region (CR) was PCR amplified using primers (TrachurusCRF2: 5'-CCAGTCCAACCTGCAAAGGA-3' & TrachurusCRR2: 5'-CACAGGGAGGCGGATACTTG-3') designed during this study. Individual PCRs were performed in 20µl total volumes, containing 50-100 ng template DNA, 2pmol of each primer,10µl of 2XBioMix (Bioline,UK) and 2µl of ddH2O. The PCR thermoprofile was 3 minutes at 95°C, followed by 35 cycles of 95°C for 30 seconds, 45 seconds annealing at 54°C for COI or 52°C for CR amplification and 1 minute at 72°C, followed by a final extension stage at 72°C for 3 minutes. Amplicons were purified and sequenced with respective forward and reverse primers using BIG DYE technology. Sequence identity was confirmed by BLAST (standard database-nucleotide collection (nr/nt)) and sequences were aligned using

the CLUSTALW program (Thompson *et al.,* 1994), implemented in BIOEDIT (Hall, 1999).

Phylogenetic relationships among haplotypes were inferred using (i) a minimum spanning network constructed in NETWORK (www.fluxus-engineering.com/sharenet.htm) and (ii) both Maximum Likelihood (ML) and Bayesian Inference (BI) trees constructed in PHYML (Guindon *et al.,* 2010) and MrBayes (Ronquist & Huelsenbeck, 2003), respectively. ML and BI trees were generated using the optimal substitution model (HKY+G) identified by ModelTest (Posada & Crandall, 1998). BI trees were run using 3 heated and one cold chain over 5,000,000 generations, with the Markov chain sampled at 1000 generation intervals, and the first 15% of trees discarded as burn-in. BI posterior probabilities were calculated over the 5,000,000 generation run and ML support was calculated based on measures of bootstrap (BS), inferred after 1000 replicates.

ARLEQUINv3.5.2.2 (Excoffier & Lischer, 2010) was used to estimate indices of haplotype diversity (*h*) and to quantify differentiation between pairs of samples, using pairwise ΦST, with significances assessed by 10,000 permutations. Fu's *F*s (Fu, 1997) and Tajima's *D* (Tajima, 1989) tests were used to test for deviations from mutation-drift equilibrium that could be attributed to selection and/or population size changes. Mismatch distributions, the frequency distributions of pairwise differences between haplotypes within a sample and simulated distributions under a model of demographic expansion, were compared using the sum of squared deviations (SSD) as a test statistic with significance assessed after 10,000 bootstrap replications. The timing of expansions (T) was estimated using the formula $T = \tau/2u$ (Rogers & Harpending, 1992). IMa2 (Hey, 2009) was used to estimate divergence times between groups with analyses being run for 1,000,000 burn-in generations and > 5,000,000 sampling generations so that the minimum ESS across parameters was > 50 (Hey & Nielsen, 2004). Mismatch and IMa2 analyses were performed using both a universal divergence rate for marine fish species mtDNA of 1.5% per million years (Bermingham *et al.,* 1997) and a fossil calibrated *Trachurus* specific divergence rate of 0.14% per million years proposed by Cardenas *et al.* (2005).

4.2.3 Microsatellite amplification & analysis

Individuals were analysed at 10 microsatellite loci 4 of which (Tt29, Tt48, Tt62, Tt113) were developed for *T. trachurus* (Kasapidis & Magoulas, 2008) and 6 (TmurA101, TmurA104, TmurA115, TmurB104, TmurB116, TmurC4) for *T. murphyi* (Canales-Aguirre *et al.,* 2010). Loci were individually PCR amplified in 10μl reaction volumes containing 50-100ng template DNA, 1pmol of each primer, 5μl of 2XBioMix (Bioline,UK) and 1μl of ddH2O. The PCR thermoprofile for Tt29, Tt48, Tt62 and Tt113 amplification was: 95°C for 3 minutes, followed by 45 cycles of 95°C for 30 seconds, 30 seconds annealing at 54°C (for Tt29, Tt48, Tt113) or 52°C (for Tt62),

and a 30 seconds extension step at 72°C, followed by a final extension step of 72°C for 3 minutes. The other 6 microsatellite loci (TmurA101, TmurA104, TmurA115, TmurB104, TmurB116, TmurC4) were PCR amplified with the following thermoprofile: 94°C for 3minutes, followed by 35 cycles of 94°C for 40 seconds, 30 seconds annealing at 57°C (TmurA104, TmurA115, TmurB104, TmurB116, TmurC4) or 52°C (TmurA101) and 30s at 72°C, with a final extension step of 72°C for 5 minutes. Amplicons were separated on an Applied Biosystems 3730 and alleles were inferred using Peak Scanner software (Applied Biosystems).

Genetic variation within samples was characterised using the number of alleles (NA), allelic richness (AR), observed heterozygosity (HO), and expected heterozygosity (HE), all calculated using GENALEX 6.2 (Peakall & Smouse, 2006). GENALEX was also used to test for genotype frequency conformance to Hardy-Weinberg expectations (HWE) and genotypic linkage equilibrium between pairs of loci. As locus x sample tests of HWE yielded a large number of deviations due to heterozygote deficits the software FreeNA (Chapuis & Estoup, 2006) was used to estimate frequencies of null alleles. Genetic differentiation among samples was quantified using *F*ST values with significance assessed following 9,999 permutations in GENALEX with pairwise *F*ST visualised using a PCoA. Global and pairwise *F*ST values were also recalculated using the null allele correction in FreeNA. Genetic structure was also investigated with/ without prior sample definition and a combination of admixture/no admixture models, using the model based Bayesian clustering programme STRUCTURE V2.3.4. (Pritchard *et al.,* 2000). Simulations were run over 500,000 MCMC iterations with the first 100,000 repetitions dis-

Table 4.1: Genetic diversity across sampling location based on: A) 1075bp of concatenated COI & CR. N= number of samples, H= number of haplotypes, *h*= haplotype diversity, π = nucleotide diversity. Standard deviation given in parentheses. B) 10 microsatellite loci. N= sample size, N_A =mean number of alleles across loci, AR= Allelic richness, H_O=mean observed heterozygosity, H_E= mean expected heterozygosity, *p*= probability of deviation form Hardy-Weinburg expectations Standard error given in parentheses. Significance indicated in bold.

	A) mtDNA (concatenated COI & CR)				B) Microsatellite					
	N	**H**	***h***	**π**	**N**	**N_A**	**AR**	**H_O**	**H_E**	***p***
Morocco: Al Hoceima	15	9	0.876 (0.070)	0.0015 (0.0003)	13	7.5	5.939 (0.526)	0.504 (0.051)	0.731 (0.060)	**<0.001**
Portugal: Olhao	16	14	0.983 (0.028)	0.0022 (0.0003)	12	6.8	5.542 (0.471)	0.485 (0.071)	0.695 (0.059)	**<0.001**
Morocco: Essaouira	22	16	0.931 (0.046)	0.0019 (0.0003)	48	10.8	6.193 (0.515)	0.539 (0.058)	0.766 (0.053)	**<0.001**
Morocco: Tiznit	21	14	0.895 (0.061)	0.0022 (0.0004)	35	10.7	0.611 (0.472)	0.544 (0.044)	0.758 (0.046)	**<0.001**
Ghana: Tema	20	14	0.916 (0.055)	0.0027 (0.0005)	49	11.2	6.297 (0.412)	0.566 (0.031)	0.781 (0.037)	**<0.001**
Angola: Namibe	9	9	1.000 (0.052)	0.0026 (0.0005)	9	6.3	5.650 (0.595)	0.691 (0.070)	0.681 (0.055)	**0.138**
South Africa: Saldanha	16	8	0.758 (0.110)	0.0014 (0.0003)	23	10.1	6.627 (0.353)	0.619 (0.048)	0.794 (0.030)	**<0.001**

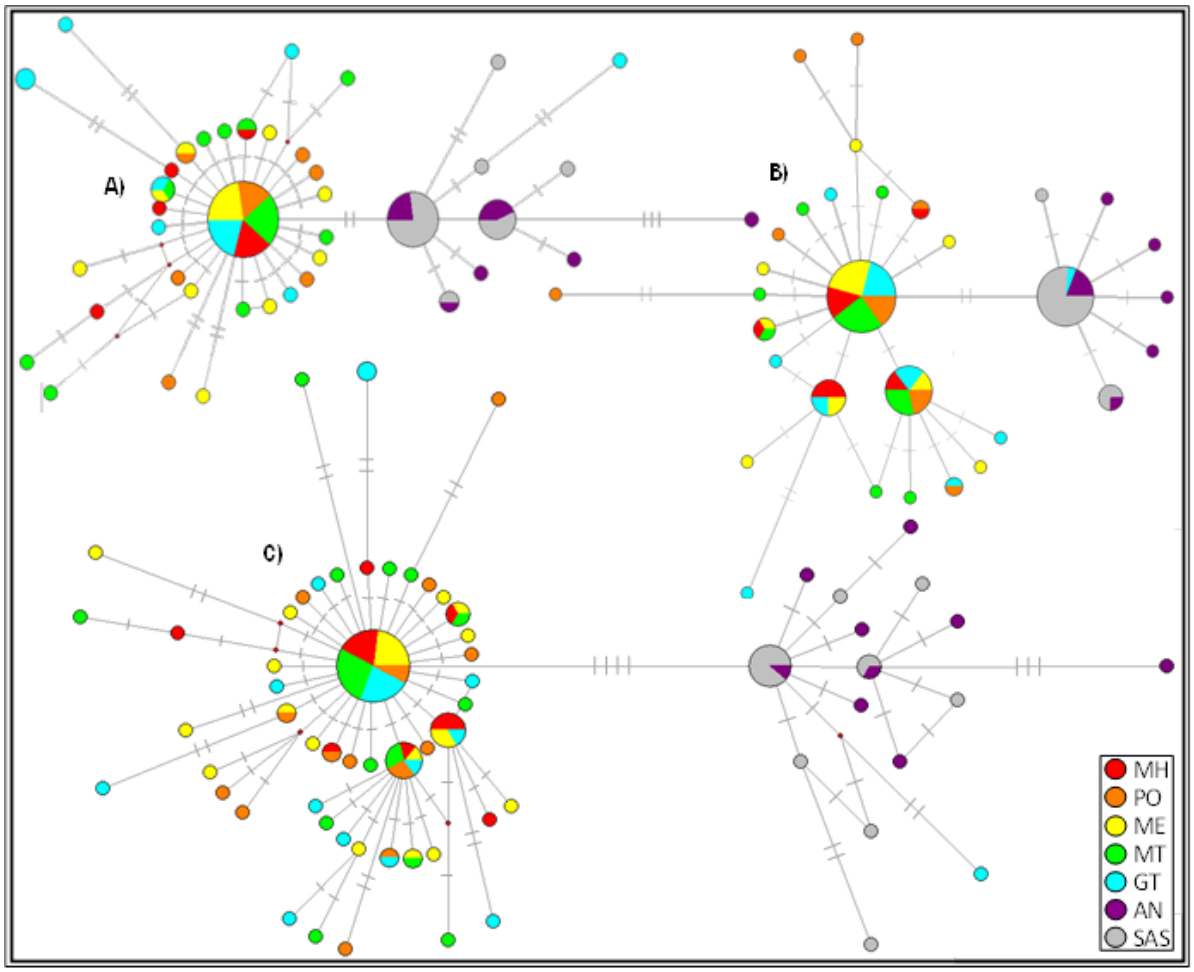


Figure 4.2: Median joining haplotype networks of T. trachurus and T. capensis based on: A) 605bp of mtDNA COI, B) 470bp of mtDNA CR, and C) 1075bp of concatenated sequence. Each dash represent a single site difference, node sizes are proportional to the overall haplotype frequency and pie discs correspond to within sample frequencies.

carded as burn-in. Optimal *K* values [from a range of 1-5] were identified using the log probability of data [Pr(X/K)]. To complement the STRUCTURE analyses classical assignment tests were performed in GENALEX. Self-assignment tests using the leave one out method were first used to assess the power of the analysis and then assignment of individuals treated as of unknown origin were performed as directed by the other results.

4.3 Results

4.3.1 MtDNA sequence data

A 605bp fragment of COI (n = 126 individuals) and 470bp fragment of the CR (n = 125 individuals) region were aligned. Levels of diversity were extremely similar for both the COI and CR. The COI alignment revealed 46 polymorphic sites defining 41 haplotypes while for CR there were 33 variable sites and 31 haplotypes. Concatenated COI and CR sequences defined 66 haplotypes. The 66 haplotypes separated into two shallowly diverged reciprocally monophyletic clades (Figure 4.2 & 4.3). There was a clear spatial separation of these clades with both

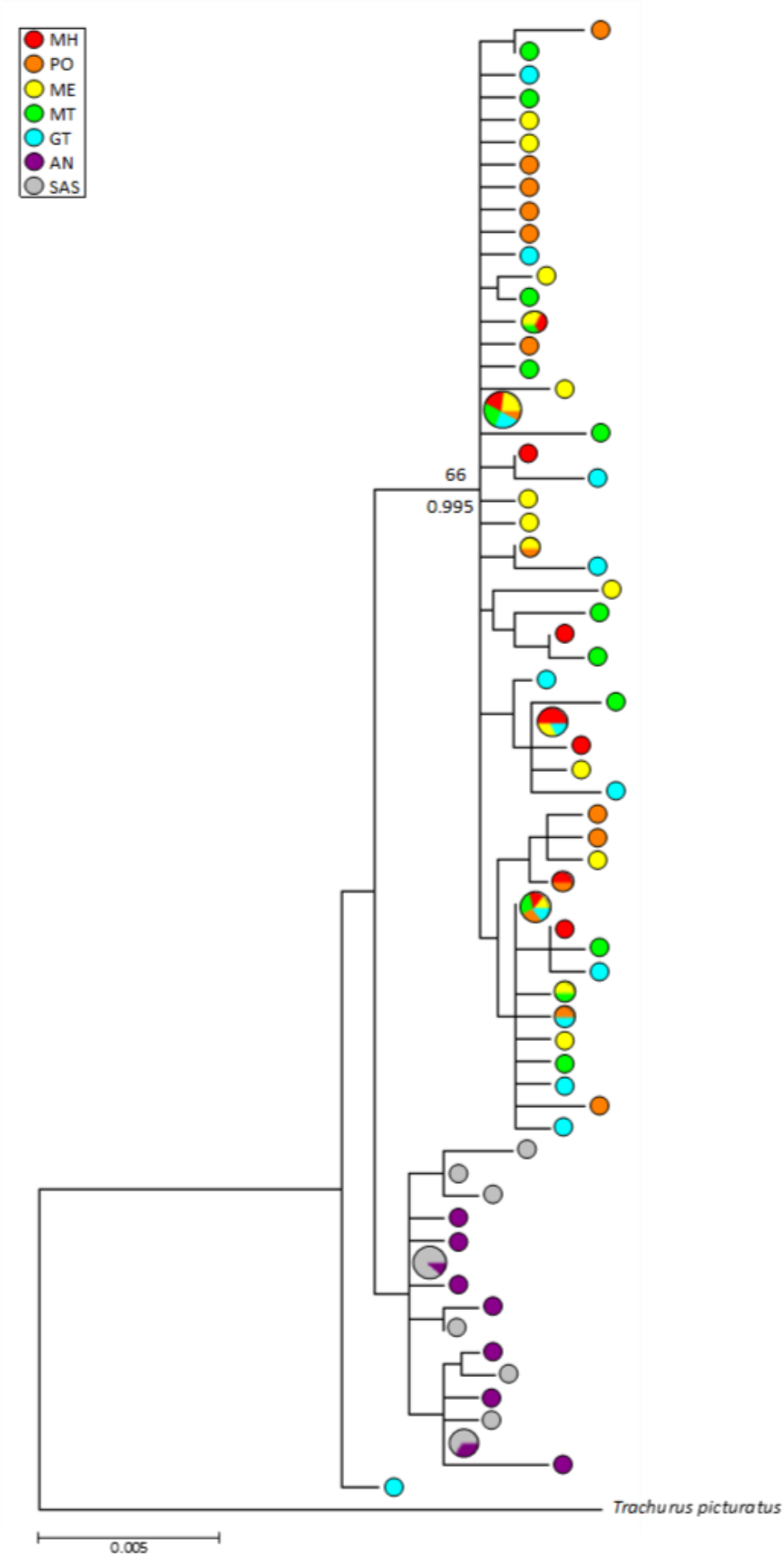


Figure 4.3 Maximum likelihood phylogeny of *T. trachurus* and *T. capensis* haplotypes resolved from 1075bp of concatenated mtDNA COI & CR with bootstrap support given above branches. Statistical support for nodes from the corresponding Bayesian analyses (BI) are reported below branches. OTU pie charts denote the proportional representation of a haplotype among sites.

Table 4.2: Pairwise mtDNA Φ_{ST} values between samples based on 1075bp of concatenated COI and CR. Significant values denoted in bold.

	MH	PO	ME	MT	GT	AN
PO	0.036					
ME	-0.024	0.012				
MT	-0.014	0.005	-0.012			
GT	-0.015	0.010	-0.005	-0.006		
AN	**0.655**	**0.619**	**0.636**	**0.610**	**0.557**	
SAS	**0.717**	**0.681**	**0.687**	**0.665**	**0.617**	-0.004

clades only co-occurring in the Ghana sample. One was found across the NE Atlantic and Ghana (and was absent from Angola and South Africa). The other clade was found only in South Africa, Angola and a single Ghanaian individual (Figure 4.2 & 4.3). These clades are hereafter referred to as the *T. trachurus* and *T. capensis* clades respectively based on BLAST. Based on the concatenated sequences mean estimates of divergence time between the clades were 1,076,715.93 (95% probability range 614,617.94 - 1,524,916.94) and 100,493.49 (95% probability range 57,364.34 -142,325.58) years BP assuming 0.14% and 1.5% divergence rates, respectively.

Levels of intra-sample variation were similar among all samples as reported in Table 1. Pairwise tests (Table 4.2) indicated no significant differentiation within either species (only Ghana *T. trachurus* included here). Tajima's D and Fu's Fs tests for neutrality were significant and negative for both clades. Mismatch distribution patterns conformed to demographic expansion models in all cases (COI; CR; concatenated sequence) except the CR based tests for *T. trachurus*. For concatenated sequences *T. capensis* $\tau = 0.615$, leading to expansion time estimates of 38, 140 years BP and 408, 638 years BP depending on the use of a 1.5% and 0.14 % sequence divergence rate, respectively. For *T. trachurus* ($\tau = 0.6$) expansion time estimates were 37, 209 years BP and 398, 671 years BP using the same divergence rates.

4.3.2 Microsatellites

In total 10 loci were screened across 189 individuals and 7 sample locations (Table 4.1). Allele sizes differed in expected multiples of the microsatellite repeat motifs. Thirty out of 70 locus by sample location comparisons deviated significantly from Hardy-Weinberg expectations

Table 4.3: Pairwise F_{ST} values based on 10 microsatellite loci with statistically significant values in bold. Sample codes as in Table 1. Above diagonal are values obtained after correction for null alleles using the ENA method. Below diagonal are values from uncorrected data.

	MH	PO	ME	MT	GT	AN	SAS
MH		-0.001	-0.001	0.002	0.005	0.008	**0.008**
PO	0.004		-0.001	0.001	**0.012**	0.003	**0.012**
ME	-0.001	-0.003		0.002	**0.005**	**0.015**	**0.012**
MT	0.006	0.005	0.002		**0.007**	**0.023**	**0.011**
GT	0.010	**0.026**	**0.011**	**0.014**		**0.017**	**0.018**
AN	0.028	0.019	**0.038**	**0.054**	**0.041**		**0.012**
SAS	0.019	0.019	**0.021**	**0.021**	**0.029**	**0.026**	

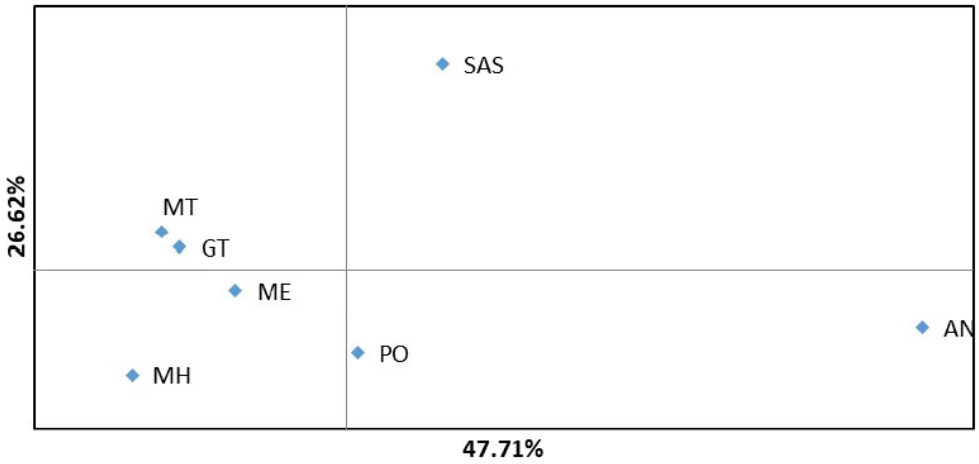


Figure 4.4: Principal coordinate analysis (PCoA) of pairwise nuclear *F*ST values estimated from the three loci showing little evidence of null alleles.

(critical *P* = 0.05). Heterozygote deficits were less prominent for three loci, with the loci Tt48, TmurB116 and TmurC4 reporting significant test results for 0, 1, and 2 sample locations, respectively (Supplementary Table 4.1). These three loci also reported the lowest average (across sample site) estimated null allele frequencies (NAF): (NAF)Tt48 = 0.008; (NAF) TmurB116 = 0.002; (NAF)TmurC4 = 0.05). Excluding these three loci the average null allele frequency per locus ranged from 0.08 (TT113) to 0.26 (A115) with a mean of 0.14 (Supplementary Table 4.2).

Based on the patterns of heterozygote deficits *FST* based analyses were performed for three data sets: (i) all loci unedited, (ii) all loci with null allele correction, and (iii) the three loci with negligible heterozygote deficits. Across all samples (both species/clades) nuclear differentiation was highly significant (*F*ST without null allele correction = 0.015; *F*ST with null allele correction = 0.016; *F*ST three loci subset = 0.015). Pairwise *F*ST values for each of the three nuclear data sets supported the differentiation of the *T. capensis* (Angolan and South Africa) from the northern samples (*T. trachurus* except for that one individual in Ghana with *T. capensis* haplotype) (Figure. 4.4). Patterns of pairwise *F*ST among the northern samples (i.e. Portugal to Ghana) differed slightly depending on the nuclear data set. For the three loci data set no significant pairwise *F*ST values were obtained (Supplementary Table 3). For the l0 loci data set with no null allele correction, the Ghana sample yielded significant *F*ST values in all pairwise comparisons, whereas after null allele correction significant *F*ST values were only obtained in comparisons between the Ghana sample and the Portugal and MT samples (Table 3).

STRUCTURE analyses performed for the unedited and three loci data sets all failed to resolve any structure with K = 1 unanimously supported. The lowest *P* for K =1 was 0.89 for the analy-

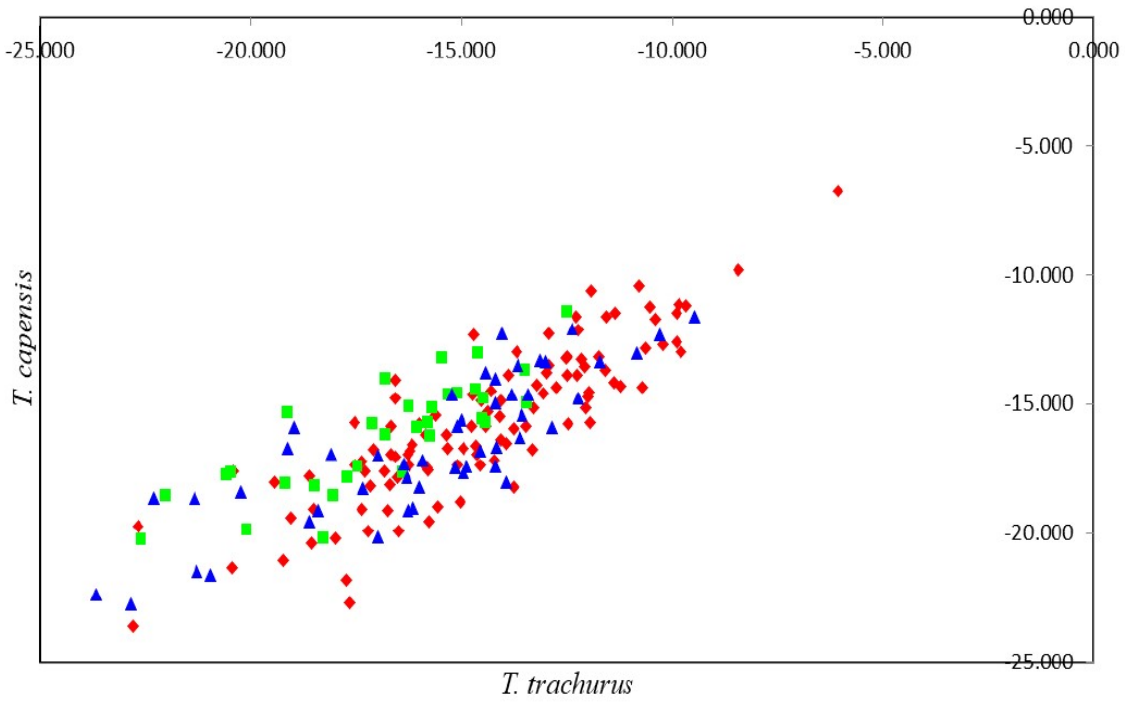


Figure 4.5: Biplot of individual self-assignment indices for the two "pure" *T. trachurus* (red) and *T. capensis* (blue) baseline groups and for the Ghana individuals (green) that were treated as unknown in the second round of assignment analysis. Lack of separation of the indices for individuals from different groups shows the limited power of the analysis.

sis performed with the three loci subset and using LocPrior and assuming no admixture. Classical assignment tests differ from STRUCTURE in that they require some *a prior* classification of baseline samples for individuals to be assigned to. We were particularly interested in assessing whether the *T. capensis* mitotype bearing individual assigned to ether *T. trachurus* or *T. capensis*. To assess the power of the analysis we first performed species self-classification tests wherein we set up a "pure" *T. trachurus* baseline population by pooling PO, MH, ME and MT samples and a "pure" *T. capensis* population by pooling the AN and SAS samples. For the unedited data set 78% of individuals correctly assigned to species (80% *T. trachurus* and 68% *T. capensis* correctly assigned). For the three loci data set the correct species assignment was lower (66% overall; 69% *T. trachurus* and 53% *T. capensis* correctly assigned). However, inspection of assignment scores indicates that even correctly assigned individuals in many cases had similar assignment scores to the other group indicating a low overall power in the analysis (Figure 5). This lack of power was epitomised in the results for the Ghana sample with the *T. capensis* haplotype. In the assignment analysis with the unedited data it assigned to *T. trachurus* with a score of 16.97 with a only slightly less likely membership score for the *T. capensis* group (19.98). With the three loci data set it also weakly assigned to *T. trachurus* (score = -2.79) but with a similar score for *T. capensis* (-3.21).

4.4 Discussion

4.4.1 Interspecific Divergence

A salient feature of the results was the resolution of two reciprocally monophyletic clades that exhibited a clear spatial partitioning. One clade was found among the NE Atlantic, Mediterranean and Ghana samples and was absent from Angola and South Africa and is hereafter referred to as the *T. trachurus* clade. The other clade was found only in South Africa, Angola and a single Ghanaian individual and is hereafter referred to as the *T. capensis* clade. BLAST results support these clade designations. Ghana was the only location where both clades were detected. There was no significant spatial structuring within either clade. Microsatellite genotypes were analysed to provide a nuclear perspective and revealed a large number of heterozygote deficits. Strong heterozygote deficits have been reported for many fish species and can be due to biological factors, such as selection, inbreeding and Wahlund effects and/or technical artefacts, most commonly null alleles (Hoarau *et al*., 2002). The association of the heterozygote deficits with certain loci and not others in this study would suggest a prominent role for null alleles. In their population genetic study of *T. trachurus* Kasapidis & Magoulas (2008) reported signals compatible with spatially variable null allele effects at three out of the four loci developed from *T. trachurus* and employed here. This study also used six microsatellite loci developed from *T. murphyi* sequences and such cross species amplification increases the likelihood of null alleles (e.g. Panova *et al*., 2006). To account for possible errors stemming from null alleles additional analyses were performed for null allele corrected genotypes and omitting loci with null alleles. A replicated pattern in *FST* based analyses across the data sets was the separation of the Angola and South Africa samples from the remaining samples. This aligns with the distribution of the *T. capensis* and *T. trachurus* clades and supports the broad scale nuclear differentiation between the species.

The shallow divergence between the *T. trachurus* and *T. capensis* mtDNA clades fits with an overall pattern of low levels of interspecific genetic divergence within *Trachurus* as reported by Cardenas *et al*., (2005). Such low levels of DNA divergence have been linked to a combination of recent divergence and low nucleotide substitution rates. Both Bektas & Belduz (2008) and Cardenas *et al*., (2005) also suggested similar low mutation rates for both coding and non-coding mtDNA regions a feature evident in the near identical levels of variation at both COI and CR here. Estimates of divergence between the *T. trachurus* and *T. capensis* mtDNA clades fitted within the broad timescales inferred by Cardenas *et al*., (2005) who suggested the divergence occurred sometime after 2.8 or 2.3 MYA depending on the divergence rate used. While the IMA based method here has been shown to perform well in cases of shallow divergences (McKeown *et al.,* 2019) a fundamental consideration for any such estimates is that commonly

inferred and even fossil obtained "species" level mutation rates may be considerably lower than those experienced at intraspecific levels and/or over more recent evolutionary time frames (1-2 MYA). This can lead to an over-estimation of the ages of divergence and demographic events (Grant, 2015; Ho *et al.,* 2005, 2007; Hoareau, 2015). Such concerns have led investigators to apply 2-10 fold rate corrections (Gonzalez-Wevar *et al.,* 2011; McKeown *et al.,* 2019; Pardo-Gandarillas *et al.,* 2018) to estimates of demographic events. While mindful of both general and *Trachurus* specific rate considerations the data are compatible with divergence between of *T. trachurus* and *T. capensis* clades during the Pleistocene, as has been observed for a number of other marine taxa (Burridge, 2002; Grant *et al.* 2005; Ludt *et al.,* 2015; Stepien & Rosenblat 1996). Applying similar rate adjustments to the dating of expansion events indicated by mismatch analysis and neutrality tests suggest that such expansions have occurred within the postglacial period, a time-frame which fits with the general view put forward by Grant (2015). In conclusion, divergence of the clades occurred during the glacial period but both lineages exhibit signals of postglacial expansions.

4.4.2 Secondary contact in Ghana

An interesting finding was the co-occurrence of both mtDNA clades in the Ghana sample with a single individual bearing the *T. capensis* clade. Secondary contact between divergent mtDNA lineages occurs in a number of fish taxa and individual assignment-based analyses have resolved cases where such secondary contact is associated with hybridisation (e.g. Sala-Bozano *et al.,* 2009) or mechanical mixing without interbreeding (Healey *et al.,* 2018a). In classical assignment tests the Ghana individual bearing the *T. capensis* clade was found to assign to the *T. trachurus* group suggesting this individual reflects recent or ancestral gene flow between the species rather than a migrant. However, an important consideration is that the low overall level of correct self-assignment to species, as well as the small differential in assignment score of many individuals to either species, suggest the data conferred little power to distinguish between a hybrid and uninformative genotype. More powerful genomic assays are needed to robustly distinguish between mixing and hybridisation. If there is a high level of introgression between both species within the region the Ghana sample would be expected to be genetically intermediate. However, PCoA of pairwise *F*ST scores clearly showed it to cluster with the *T. trachurus* samples distinct from the *T. capensis* samples suggestive of low rates of hybridisation.

4.4.3 Conservation and management implications

There was no significant mtDNA differentiation between the *T. capensis* samples from Angola and South Africa. This is a striking result in light of the high level of mtDNA divergence, including examples of reciprocal monophyly, between Angolan and South African sites for a number

of marine taxa due to the BUS biogeographic boundary (Gwilliam *et al*., 2018; Henriques *et al*., 2012; 2014; Reid *et al*., 2016). While the previously mentioned slow mtDNA mutation rates for *Trachurus* could limit differentiation effective gene flow across the BUS aligns with expectations from the species' ecology. Specifically, while the cool upwelled waters of the BUS are a dispersal barrier for many species *T. capensis* is an upwelling associated species with peak abundance observed across the main upwelling regions and is expected to freely disperse across the BUS (Axelsen *et al*., 2004; Barange *et al*., 1998; McLaverty, 2012). Microsatellites did reveal a significant *FST* value between the Angolan and SAF samples. However, in light of the small sample size for Angola and prevalence of null alleles the biological validity of this differentiation should be considered questionable unless confirmed by other studies (Knutsen *et al.,* 2011). Overall, the strongest elements of the data and life history expectations suggest connectivity between Angolan and South African *T. capensis*. The current management of the species as distinct northern and southern BUS stocks therefore represents a potential misalignment between management and population boundaries (Reiss *et al*. 2009).

MtDNA and microsatellites revealed no significant genetic structuring among the NE Atlantic and Mediterranean *T. trachurus* samples. This aligns with the previous *T. trachurus* study by Kasapidis & Magoulas (2008) which found no significant structure across a more extensive sampling of NE Atlantic and Mediterranean waters using four of the microsatellites employed here. Genetic cohesion between Atlantic and Mediterranean samples is also described for other horse mackerel species and indicates that the Almeria-Oran frontal system is not a prominent phylogeographic break in this group (Comesana *et al.*, 2008; Moreira *et al*., 2019). Nuclear *F*ST tests performed for both the null allele corrected and the subset of HWE conforming loci and mtDNA ΦST broadly supported a lack of differentiation of the Ghana sample from the other *T. trachurus* samples. Although there was a gap in our sampling between Ghana and Europe, high gene flow between European and NW African sites is also suggested for the blue jack mackerel (*T. picturatus*) (Moreira *et al.,* 2019). The available data therefore indicate that *T. trachurus* exhibits high gene flow and an absence of genetic structure across most of its sampled range to date.

An important consideration previously highlighted by Kasapidis & Magoulas (2008) is that the absence of genetic structure in *T. trachurus* and *T. capensis* does not mean there is no significant isolation of stocks on timescales of interest to management (Hauser & Carvalho, 2008). Resolution of such stock structure will require more powerful genomic approaches, such as RADseq, which offer the potential to genotype large numbers of loci and identify markers un-

der selection (Gagnaire *et al.,* 2015; Mullins *et al.,* 2018). Combining genomic approaches with information from ontogenetic markers could also provide synergistic insights. Ontogenetic markers have already proven highly informative in mackerel with studies of parasite assemblages (MacKenzie *et al*., 2008), morphometrics (Murta *et al*., 2008; Stransky *et al*., 2008) and life history traits (Abaunza *et al*., 2008b) reporting finer scale population heterogeneity than indicated by population genetic methods to date. Despite signatures of historical population size fluctuations all samples exhibited similar levels of genetic variation that were comparable to levels in other small pelagic fish species (*Sardinella aurita*: Tringali & Wilson, 1993; *Engraulis encrasicolous*: Silva *et al*., 2017; *Engraulis japonicus*: Liu *et al*. 2006; and *T. picturatus*: Moreira *et al.,* 2019). Although fisheries induced loss of genetic variation, as observed in other taxa (Cimmaruta *et al*., 2005) cannot be discounted the data indicate that local populations retain high levels of genetic variation and if current effective population sizes are maintained genetic drift alone will be insufficient to erode variation.

4.4.4 Conclusions

The results of this study both inform conservation management strategies and direct future research. The data support the view that *T. trachurus* is not found in Angolan and South African waters as has sometimes been suggested. The co-occurrence of both species' clades in Ghana also reveals this area to be an apparent suture zone between their ranges. Further sampling of the region is required to assess the extent of spatial overlap between both species. The results here suggest both species are likely being indiscriminately harvested within this region, which may severely compromise stock sustainability (McKeown *et al.,* 2015; Healey *et al.,* 2018a&b). More powerful genomic based assays will also be useful to permit individual based analyses of hybridisation and fish traceability (Hemmer-Hansen *et al.,* 2019; Helyar *et al.,* 2011). Such genomic assays will also be necessary for both species to align spatial management units with patterns of stock recruitment. Until such information is available the current designation of regional management units in *T. trachurus* should be maintained as a precautionary measure. Paradoxically, for *T. capensis* the likelihood of connectivity between the northern and southern BUS operational stocks should be considered in attempts to understand potentially cohesive stock dynamics (Frisk *et al.,* 2008).

CHAPTER 5

Evolutionary history of south-west Indian Ocean Emperor fish

Published in: Ecology and Evolution (2018)
DOI: 10.1002/ece3.3775

CHAPTER 5.1

Cryptic species and parallel genetic structuring in Lethrinid fish (*Lethrinus mahsena* & *Lethrnius harak*): Implications for conservation and management in the south-west Indian Ocean

5.1.1 Introduction

Understanding the evolution of marine biodiversity requires knowledge of the patterns and processes underpinning genetic structuring of species. Combined demographic-genetic studies can reveal the spatial and temporal scale at which evolutionary forces are occurring (Waples 1998) and, by extension, identify cryptic components of biodiversity and so allow optimisation of spatial conservation strategies. Identification of spatial and temporal processes is particularly important for harvested taxa, as genetic diversity and adaptation are key factors underpinning the resilience/sustainability, and ultimately the evolutionary potential, of populations and species (Iles & Sinclair 1982; Ryman et al. 1995; Ruzzante et al. 2006; Therkildsen et al. 2013).

The southwest Indian Ocean (SWIO), which forms a subdivision of the tropical Indo-Pacific, is renowned for its diverse marine habitats and resources, including an ichthyofaunal richness (Smith *et al.* 2003) that supports significant artisanal, subsistence and commercial fisheries across the region (Berg *et al.* 2002). The complex oceanographic features of this region, characterised by the bifurcation of the South Equatorial Current into two continental currents and the series of eddies in the Mozambique Channel (Benny 2002), are hypothesised to have contributed to this high biodiversity (Ridgway & Sampayo 2005). Diverse patterns of genetic differentiation have been observed within species distributed across the SWIO, from high levels of gene flow (Duda & Palumbi 1999; Chiang *et al.* 2008; Muths *et al*. 2009; 2013; Fratini *et al*. 2010; Ragionieri *et al*. 2010; Silva *et al*. 2010b) to more complex species-specific patterns of isolation, including disruption of gene flow across the Mozambique Channel (Gopal *et al.* 2006; Silva *et al.* 2010a; Visram *et al.* 2010; Muths *et al.* 2011) and isolation of the northern (Ragioneiri *et al.* 2010; Visram *et al.* 2010) and/or southern (Fratini & Vaninni 2002; Lessios *et al.* 2001; Muths *et al.* 2011; 2015) Mascarene Islands from continental populations. Some investigations have found signatures of historical vicariance associated with sea level fluctuations during the Pleistocene (Ragionieri *et al.* 2009; Silva *et al.* 2010a). However, despite a growing body of ecological and evolutionary research the SWIO is still regarded as one of the least-studied tropical ecosystems (Ridgway & Sampayo 2005; Beheregaray 2008).

The complex array of genetic population structures resolved to date within the SWIO highlight the potential hazards of making region-wide predictions of population structure from single species. However, studies of ecologically similar and contemporaneously co-distributed species represent analytical frameworks wherein variance between lineages, place and time are minimised, and thus 'general rules' informative to subsets of taxa may be resolved (e.g. Bohonak 1999; Lester *et al*. 2007;Bird *et al.* 2007; 2011; Dawson 2012). In this study, we develop such an analytical framework by performing replicated genetic analysis of two lethrinid fish species: *Lethrinus harak*, the thumbprint emperor and *Lethrinus mahsena*, the sky emperor. Both species are ecologically and biogeographically similar. The habitat of *L. harak*, the most abundant and commercially important lethrinid in the SWIO (Kulmiye *et al.*2002), consists of shallow and protected coastal areas including reefs, mangroves, shallow lagoons and seagrass beds (Kulmiye *et al.* 2002; Ebisawa & Ozawa 2009). *Lethrinus mahsena* occupies similar shallow habitats but has been observed predominantly in reefs or adjacent areas (Carpenter & Allen 1989). Like the majority of lethrinids, both species are relatively long-lived (up to 15 years) (Ebisawa & Ozawa 2009) protogynous hermaphrodites (Ebisawa 2006; Grandcourt 2002) but many aspects of their reproductive biology are not known (Kulmiye *et al.*2002; Ebisawa 2006). Both species have pelagic larval stages with Nakamura *et al*., (2010) reporting a larval duration of around 29 days for *L. harak*.

The aim was to test the prediction that both species would display similar patterns of historical (phylogeographic) and contemporary population genetic structure across the SWIO. As the Lethrinidae is one of the most important families of fish for artisanal, recreational and subsistence fisheries (Carpenter & Allen 1989; Gouws 2012), the genetic patterns are interpreted in the context of recurrent recruitment processes relevant to stock sustainability. To this end, kinship analyses were employed to complement estimates of connectivity / isolation based on traditional population genetic (e.g. F_{ST}) methods. To provide an additional historical context, we combined genetic studies with ecological niche modelling (ENM), which has been shown to provide novel insights into the influence of historical biogeography on genetic variation in marine species (Glynn *et al.* 2015; 2016).

5.1.2 Methods

5.1.2.1 Sample collection and molecular analyses

Samples (fin clips fixed in 95% ethanol) were collected from commercial or subsistence catches landed in seven locations across the SWIO between 2009 and 2012 (Fig. 5.1.1). Total DNA was extracted from all samples using a standard CTAB-chloroform / isoamylalcohol method (Winnepenninckx *et al.*1993).

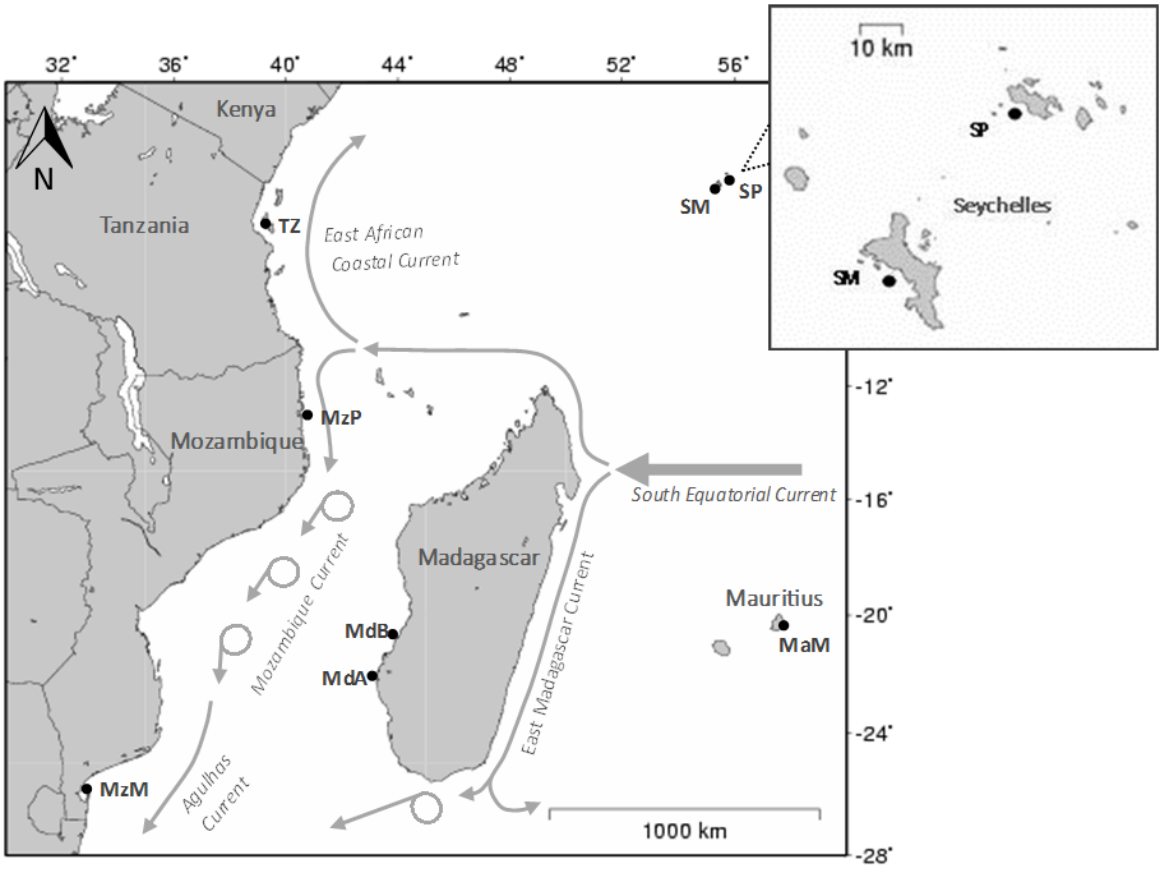


Figure 5.1.1: Sampling locations for *Lethrinus mahsena* and *Lethrinus harak* within the South West Indian Ocean: Mauritius, Mahebourg (MaM); Seychelles, Praslin (SP); Seychelles, Mahe (SM); Tanzania, Zanzibar (TZ); Mozambique, Pemba (MzP); Madagascar, Belo sur Mer (MdB); Mozambique, Maputo (MzM).

For both species (across a subset of individuals) a 486bp fragment of the mtDNA Cytochrome *c* oxidase subunit I (COI) gene was amplified by polymerase chain reaction (PCR) with species-specific primers (*L. mahsena:* LmCOIF 5'-TGGTAGGAACAGCCCTAAGC-3' & LmCOIR 5'- AGAATT-GGGTCCCCTCCTC-3'; *L. harak:* LhCOIF 5' –CGAACTTAGTCAGCCCGGA-3' & LhCOIR 5'- TGCTGA-TAGAGGATTGGGTC-3') for a subset of individuals. PCRs were performed in a total volume of 20µl, containing 4 µL template DNA, 2 mM $MgCl_2$, 0.5 µM forward primer and 0.5 µM of reverse primer, 0.2 mM dNTP mix (20 µM each dATP, dCTP, dGTP, dTTP), 1x reaction buffer [75 mM Tris-HCl, 20 mM $(NH_4)_2SO_4$] and *Taq* polymerase (BIOTAQ, 5 U/µl). The PCR thermoprofile was: 180s at 95°C, followed by 45 cycles of 30s at 95°C, 45 seconds annealing at 55°C for both primer sets and 60s at 72°C, followed by a final 5 minute extension at 72°C. PCR products were then purified using EXOSAPIT and sequenced from both directions on an Applied Biosystems 3500 platform using the respective PCR primers. Sequences were aligned using the CLUSTAL W (Thompson *et al.*1994) programme, available in BIOEDIT (Hall 1999), and analysed using BLAST. Sequences revealed the presence of a highly divergent clade among the Seychelles *L. mahsena* samples (described in results). To investigate the frequency of occurrence of this

clade elsewhere a diagnostic PCR-RFLP method was developed and used to genotype all *L. mahsena* samples (Appendix: Fig. A1).

Nuclear genetic variation was assessed at nine dinucleotide (100RTE, 95ACRTE, 95TGRTE, 90RTE, 80RTE, 75RTE,68RTE, 58RTE, BST2.33) and one trinucleotide (96RTE) microsatellite loci described by Van Herwerden *et al.*(2000; 2003). PCRs were performed under the following conditions: 180s at 95°C, followed by 35 cycles of 30 seconds denaturing at 95°C, 30 seconds annealing at 50°C and 30 seconds at 72°C. In order to increase product strength and reduce non-specific products, PCR protocols were altered for a subset of loci. Annealing temperature was increased to 60°C for 96RTE and 95ACRTE. For 75RTE, 68RTE and 80RTE the number of cycles was increased to 55 and annealing temperatures increased to 60°C, 55°C and 54°C, respectively. Each 10µl reaction contained 3µL template DNA, 1pmol of each primer, 5µL of 2xBioMix [1.5mM MgCL2] (Bioline, UK) and 1µL of ddH_2O. Amplicons were separated on an Applied Biosystems 3730 with alleles inferred using the PEAKSCANNER software (Applied Biosystems). To ensure robustness of genotypes we employed a double genotyping as described by McKeown *et al.* (2017a).

2.2 Statistical analysis of mtDNA sequences

After pruning sequences, a total of 486bp of the mtDNA COI gene was aligned across 72 *L. mahsena* and 104 *L. harak* samples. All analyses of mtDNA sequences were conducted in Arlequin 3.5.1.2 (Excoffier & Lischer 2010) unless otherwise stated. Phylogenetic relationships among sequences were inferred using maximum likelihood (ML) trees in MEGA v6.06 (Tamura *et al.* 2013) and Bayesian inference performed using MRBayes v3.2 (Ronquist & Huelsenbeck 2003). For both species the K2P+G+I substitution model, identified as optimal using MODELTEST 3.7 (Posada & Crandall 1998), was used. Maximum likelihood bootstrap values were calculated using 1000 bootstrap replicates and Bayesian Inference (BI) was calculated assuming unknown model parameters, and run over 5,000,000 generations, sampling the Markov chain every 1000 generations, with the first 15% of trees discarded as burn-in. Median joining haplotype networks were constructed in NETWORK 5.0 (Fluxus Technology). Percentage sequence divergence within and between species/clades and K2P distances were calculated using MEGA v6.06.

Genetic variation was assessed using calculations of the number of haplotypes (H), in addition to indices of haplotype (h) and nucleotide (π) diversity (Nei 1987) alongside their variances. Genetic differentiation among samples was further tested by global and pairwise Φ_{ST} using pairwise haplotype distances (Weir & Cockerham 1984), with associated P values estimated after 10,000 permutations.

5.1.2.3 Statistical analysis of microsatellite data

The number of alleles (NA), allelic richness (*AR*) and observed (H_O) and expected (H_E) heterozygosities were calculated using GENALEX 6 (Peakall & Smouse 2006). Deviations from Hardy-Weinberg (HW) expectations and linkage disequilibrium between pairs of loci were assessed using exact tests in GENEPOP v4.2 (Raymond & Rousset 1995).

Genetic differentiation was quantified using global and pairwise *F*-statistics (Wright 1978). To account for possible null allele effects such indices were also calculated using the null alleles adjustment in FreeNA (Chapuis & Estoup 2007). To test for signatures of Isolation by Distance (IBD), correlations between geographic distances (minimum sea distance in nautical miles) and linearly transformed genetic differences (F_{ST}/(1- F_{ST}) Slatkin 1993) were tested using a MANTEL matrix correlation test in GENALEX. Genetic structure was also investigated using Bayesian 'group assignment' 'without admixture' methods implemented in BAPS 6.0 (Corander *et al*.2004), for models of K=1-6 (10 independent runs per K). Genetic relationships among samples were visualised using factorial correspondence analysis (FCA) in GENETIX (Belkhir et al. 2004). Additionally, to assess nuclear differentiation between descendants of the two highly divergent mtDNA clades observed in *L. mahsena* (Seychelles), self-classification tests (i.e. assignment of nuclear genotypes to clade defined baseline groups) were employed in GENECLASS 2 (Piry *et al.*2004). The Chord genetic distance (Cavalli-Sforza & Edwards 1967) was used due to the deviations from Hardy-Weinberg equilibrium.

Mean pairwise relatedness within samples was calculated using the relatedness estimator r_{qg} of Queller & Goodnight (1989) in GENALEX (Peakall & Smouse 2006) with associated 95% confidence intervals determined by 1000 bootstraps. Permutation of genotypes among all samples (999 times) was used to calculate the upper and lower 95% confidence intervals for the expected range of r_{qg} under a panmictic model. The maximum likelihood method implemented in ML-RELATE (Kalinowski et al. 2006), was used to infer the relationships among pairs of individuals, specifically to categorise them as unrelated (U), half-sib (HS) ,full-sib (FS) or parent -offspring (PO). Pairs of individuals were classified into three categories: (a) unrelated (classification as U only), (b) related (classification as any combination of HS, FS, PO but not U), or (c) ambiguous (classification as U as well as some related state).

5.1.2.4 Estimation of Type I and Type II error rates

To assess differences in statistical power among the data sets for both species and marker types, Type I and Type II error rates were estimated using POWSIM (Ryman & Palm 2006). Analyses were run for a combination of the average and smallest sample sizes for each species (adjusted for mtDNA following Larsson *et al.* 2009).

5.1.2.5 Palaeodistribution modelling

To evaluate suitable habitat ranges (and any geographical range changes) of *L. harak* and *L. mahsena* during the Last Glacial Maximum (LGM; ~21KYA), palaeodistribution modelling was conducted. Post 1950's species occurrence data were obtained from the Global Biodiversity Information Facility (GBIF; http://www.gbif.org), along with the Ocean Biogeographic Information System (http://www.iobis.org), resulting in a total of 335 occurrences of *L. harak* and 152 of *L. mahsena*. Contemporary bioclimatic data, specifically generated for marine environments, were collected at 5-min resolution from MARSPEC (Sbrocco & Barber 2013: Supplementary Table 5.1.1). Models were generated in MAXENT v3.3.3 (Phillips *et al.*2006) and cross-validation conducted using 10 replicate runs, under the default parameters. Model performance was assessed based upon the area under the received operating characteristic curve (AUC). Finally, the models were projected onto bioclimatic data reconstructed to represent the LGM (Sbrocco 2014), using an ensemble of five different models (CNRN, ECBILTCLIO, FGOALS, Had-CM and MIROC-322).

5.1.3. Results

5.1.3.1 Cryptic genetic divergence in sympatry

Phylogenetic reconstruction of COI haplotypes among putative *L. mahsena* revealed the presence of two highly divergent clades within the SWIO, which were separated by 20 mutational steps (Fig. 5.1.2). Diagnostic PCR-RFLP analysis of all samples (Appendix A) confirmed that Clade A individuals occurred across all samples (Mauritius, Mahe island Seychelles, Tanzania, Mozambique Pemba, Mozambique Maputo, Madagascar Belo sur Mer) excluding Praslin island Seychelles, whereas clade B individuals were found in the Seychelles samples, where they co-occurred with, and were more frequent than, clade A fish.

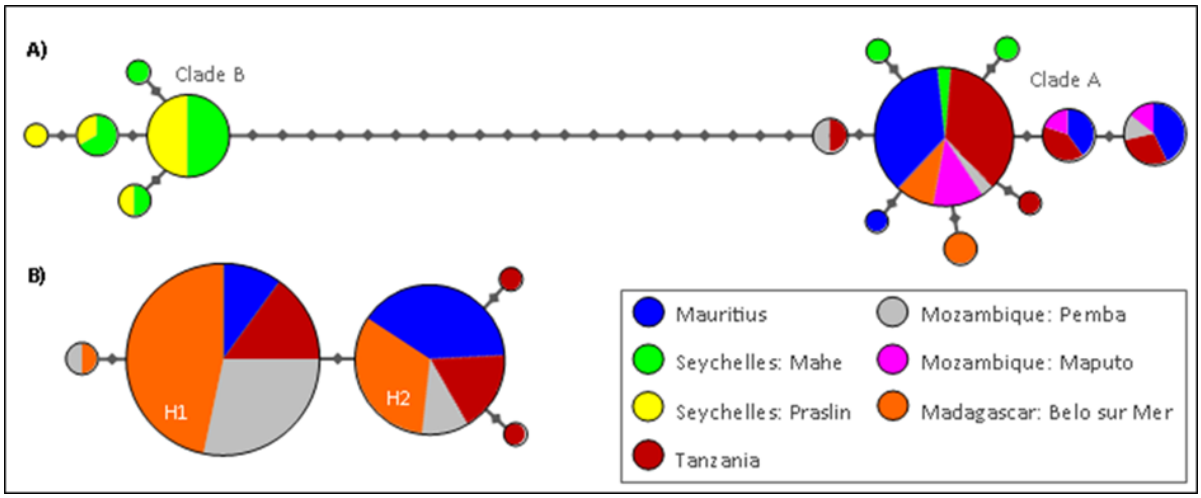


Figure 5.1.2: Haplotype network of A) *Lethrinus mahsena* and B) *Lethrinus harak* based on 486bp of COI mtDNA. Each small grey diamond represents a single mutation; branch lengths are proportional to the number of differences and node size is proportional to the haplotype frequency

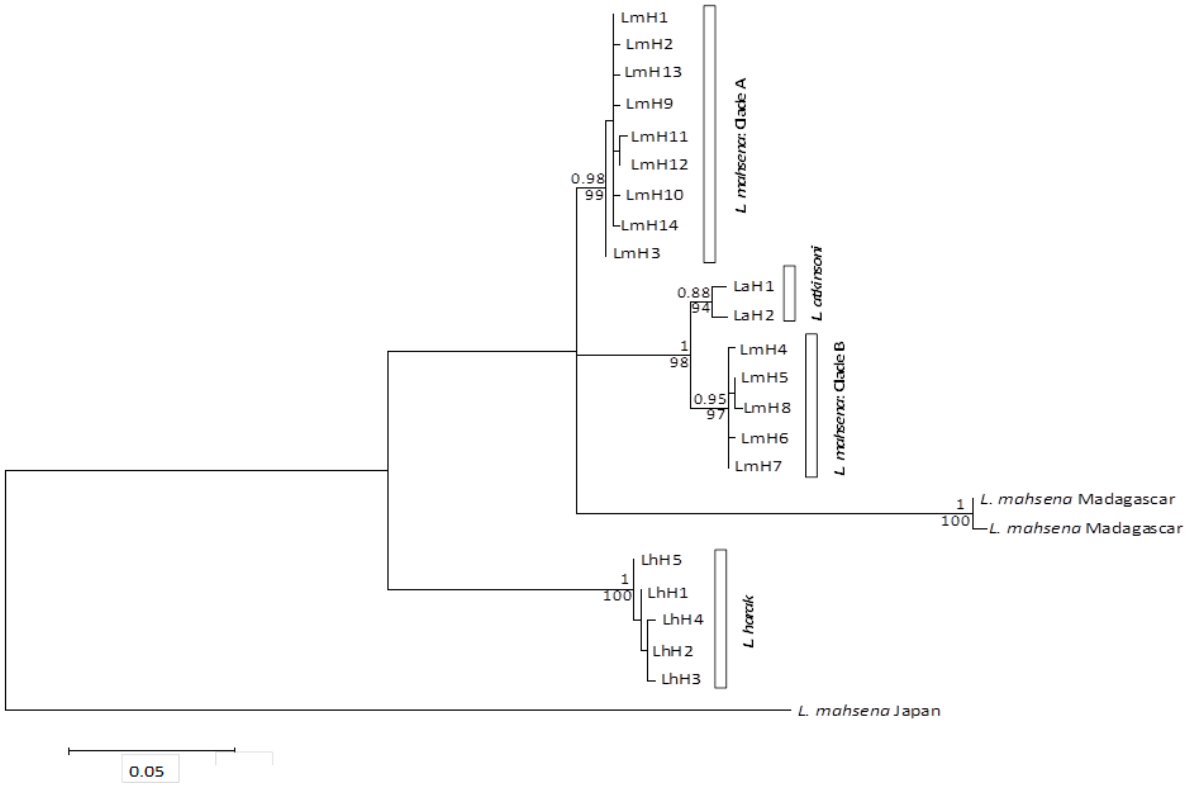


Figure 5.1.3: Phylogenetic relationships within *Lethrinus* species using 486bp COI mtDNA. Statistical support for nodes is given for both Bayesian analyses (posterior probabilities) above branches and ML analyses (bootstrap support) below branches. Branch tips are labelled as such: Lm = *Lethrinus mahsena*, La = *Lethrinus atkinsoni*, Lh = *Lethrinus harak*. Haplotype (H) numbers for *L.* mahsena and *L. harak* correspond to supplementary Tables 2 & 3, for *L. atkinsoni* H1= GENBANK sequences from Queensland, Australia (KP194639.1, KP194307.1) laH2= GENBANK sequences from the Philippines (KF009603.1, KC970391.1). *L. mahsena* Madagascar and *L. mahsena* Japan = sequences deposited as *L. mahsena* on GENBANK (JQ350088.1, JQ350089.1, JF952782.1) that do not correspond phylogenetically to this classification.

BLAST searches of clade A individuals revealed a 99% sequence similarity to *L. mahsena* voucher specimens from Mozambique (JF493750.1, JF493751.1, JF493752.1), however BLAST searches of clade B individuals also identified 99% sequence similarity to *L. mahsena* voucher specimens from India (EF609387.1, KM079305.1, KM079304.1, KJ920117.1); this was followed by 97% sequence similarity to all available *L. atkinsoni* voucher specimens from Western Australia, the Philippines and Taiwan (KV944053.1, KP194639.1, KP194152.1, KP194307.1, KF009603.1, KC970391.1, EF609384.1). The average sequence divergence between clade A and clade B was 4.6% (k2p distance=0.048, SE=0.010) (Supplementary Table 5.1.4) with clade B less divergent from *L. atkinsoni* (2.2%; k2p distance=0.021, SE=0.006) (Fig. 5.1.3). Microsatellite genotypes revealed a clear differentiation between individuals partitioned according to their mtDNA clade in the FCA (Fig. 5.1.4), with a corresponding high F_{ST} (0.01, P<0.001). This pattern was also supported by assignment tests which found complete self-classification to clade with the exception of a single individual from clade A, which had a similar probability of assignment to both clades (0.55 and 0.45 assignment probability to Clade A and B groups, respectively).

5.1.3.2 MtDNA COI sequence data

Overall haplotype and nucleotide diversities (Table 5.1.1) were similar for both *L. harak* (h=0.52, π=0.0011) and *L. mahsena* (h=0.59, π=0.0018). Network representation of *L. mahsena* (excluding clade B) revealed a single haplogroup dominated by a single common haplotype present across all SWIO localities. Similarly for *L. harak*, a single haplogroup was dominated by two common haplotypes (Fig. 5.1.2): haplotype 2 which was dominant in Mauritius and haplotype 1 which was dominant across all other SWIO localities.

Although global values of genetic differentiation were high and significant for *L. harak* (Φ_{ST}= 0.151, *P* = 0.002). pairwise tests revealed all significant values to be associated with two samples (Mauritius & Mozambique Pemba), with only the comparison between Mauritius and Mozambique Pemba remaining significant after Bonferroni correction (Table 5.1.2).Global differentiation was lower and not significant for *L. mahsena* (Φ_{ST}=0.024, *P* = 0.277) and although pairwise Φ_{ST} were high in many cases, they were all non-significant (Table 5.1.2).

Owing to small sample sizes, demographic tests were conducted globally. Neither Tajima's *D* (*D* = -0.524, *P* = 0.335) nor Fu's *Fs* (*Fs* = -1.100, *P*=0.280) tests detected global deviations from neutrality for *L. harak*. For *L. mahsena* Fu's *Fs* was significant (*Fs*=-4.617, *P*=0.002) but Tajima's *D* was not (*D* = -1.330, *P* = 0.070). Mismatch analyses were compatible with expansions for *L. mahsena* (τ = 1.078, 95% CI: 0-3.43, *P*=0.690) but not in *L. harak* (τ = 0.723, 95% CI: 0.473-1.076, *P*=0.010).

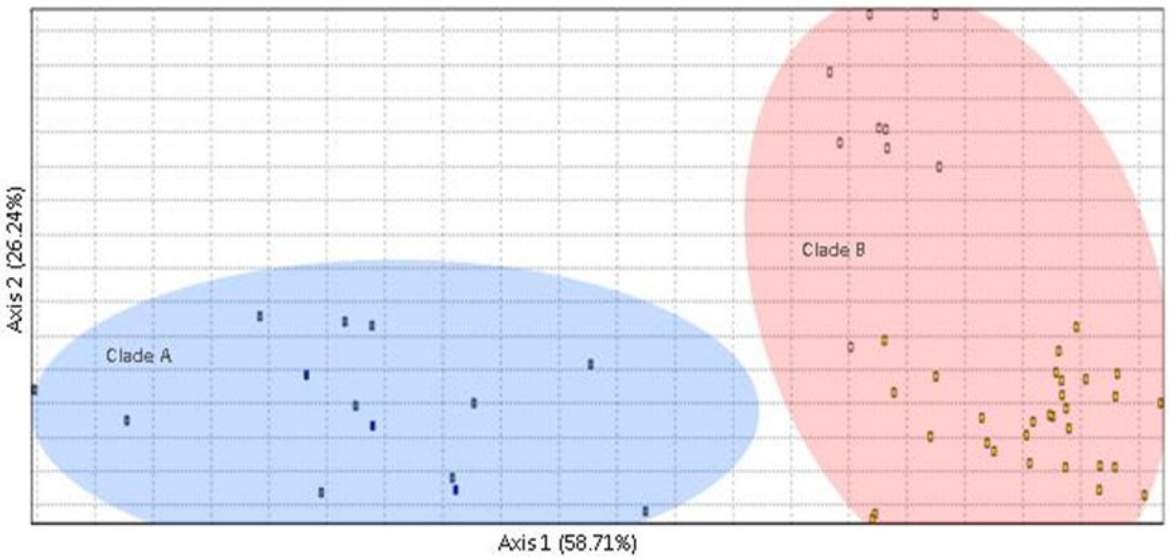


Figure 5.1.4: Factorial correspondence analysis showing multivariate relationships among microsatellite genotypes of *Lethrinus mahsena* individuals from Mahe (SM) and Praslin (SP) islands in the Seychelles. Samples were sorted by mitochondrial CO1 clade (clade A/ clade B), classified by through PCR-RFLP analysis.

Table 5.1.1: (a) Mitochondrial genetic diversity levels, mtDNA neutrality tests, and a summary of mismatch distributions for 486 bp of *Lethrinus harak* and *Lethrinus mahsena* COI mtDNA: mtDNA sample size = *N*, number of haplotypes = *H*, number of private haplotypes = pHap, haplotype diversity = *h*, nucleotide diversity = π. (b) Microsatellite genetic diversity across nine (*L. harak*) and 10 (*L. mahsena*) loci. Microsatellite sample size = N, number of alleles = NA, observed heterozygosity = H_O, expected heterozygosity = H_E, *p*= significance of Hardy-Weinberg tests.

	a) mtDNA (COI)					b) Microsatellite				
	N	H	Phap	h	π	N	N_A	H_O	H_E	*p*
Mauritius										
L. harak	22	2	0	0.42 (0.09)	0.0008 (0.0009)	43	4.00	0.52	0.45	**0.000**
L. mahsena	18	4	1	0.54 (0.12)	0.0017 (0.0014)	19	4.40	0.59	0.56	**0.023**
Seychelles: Mahe										
L. mahsena	3	3	2	1.00 (0.27)	0.0027 (0.0028)	12	3.60	0.58	0.50	**0.044**
Tanzania										
L. harak	18	4	2	0.63 (0.07)	0.0015 (0.0013)	60	4.11	0.42	0.42	**0.000**
L. mahsena	18	5	1	0.57 (0.13)	0.0016 (0.0014)	21	4.40	0.54	0.53	0.412
Mozambique: Pemba										
L. harak	22	3	0	0.39 (0.11)	0.0008 (0.0008)	44	3.89	0.41	0.42	**0.000**
L. mahsena	3	3	0	1.00 (0.27)	0.0041 (0.0039)	7	3.10	0.46	0.45	0.246
Mozambique: Maputo										
L. mahsena	6	3	0	0.60 (0.22)	0.0018 (0.0020)	10	3.60	0.45	0.49	0.053
Madagascar: Belo sur Mer										
L. harak	42	3	0	0.47 (0.06)	0.0009 (0.0009)	100	4.78	0.42	0.42	**0.000**
L. mahsena	5	2	1	0.60 (0.18)	0.0012 (0.0014)	5	2.30	0.53	0.40	0.991
Total										
L. harak	104	5	2	0.52 (0.03)	0.0011 (0.0019)	247	4.19	0.44	0.42	**0.000**
L. mahsena	53	9	7	0.59 (0.07)	0.0018 (0.0036)	72	3.57	0.52	0.49	**0.025**

5.1.3.3 Microsatellite allele frequency data

A total of 247 individuals of *L. harak* from 4 sample locations and 72 *L. mahsena* individuals (excluding those belonging to clade B) from 6 sample locations were screened across 9 and 10 microsatellite loci respectively. Levels of intra-sample genetic variability were similar in both species (Table 5.1.1). No significant linkage disequilibrium was indicated between any pair of loci in either species. POWSIM analysis indicated that microsatellite data conferred more power to detect genetic differences than mtDNA, with a similar resolution for both species (Supplementary Table 5.1.5), and with generally low Type I error probabilities.

For *L. mahsena* 8 out of 60 locus by sample tests revealed significant deviations from HWE (Supplementary Table 5.1.6), primarily due to heterozygote deficits spread relatively consistently across 5 loci (80RTE, BST2.33, 100RTE, 95TGRTE and 68RTE). For *L. harak*, 12 out of 36 tests demonstrated significant deviations from HWE (Supplementary Table 5.1.7), again largely due to heterozygote deficits with locus 58RTE exhibiting a deficit of heterozygoites in all 4 samples.

Global estimates of differentiation (with null allele correction) was higher for *L. mahsena* (F_{ST}= 0.164) than *L. harak* (F_{ST}= 0.023) but highly significant in both cases. Pairwise F_{ST} estimates were also high and significant between all samples for both *L. harak* and *L. mahsena* (Table

Table 5.1.2: Pairwise estimates of genetic differentiation (*FST*) between *Lethrinus mahsena* samples (below diagonal—across 10 microsatellite loci corrected for null alleles (msat) and 486 bp of mtDNA COI) and *Lethrinus harak* samples (above diagonal—across nine microsatellite loci corrected for null alleles (msat) and 510 bp mtDNA COI)

	Mauritius	Seychelles Mahe	Tanzania	Mozambique Pemba	Mozambique Maputo	Madagascar Belo sur Mer
Mauritius						
COI			0.049	**0.407***		**0.256**
msat			**0.039***	**0.042***		**0.029***
Seychelles: Mahe						
COI	0.152					
msat	**0.121 ***					
Tanzania						
COI	-0.038	0.127		**0.127**		0.035
msat	**0.043***	**0.241***		**0.022***		**0.012***
Mozambique: Pemba						
COI	-0.044	0.000	-0.043			0.002
msat	**0.118***	**0.128***	**0.239***			**0.013***
Mozambique: Maputo						
COI	-0.116	0.100	-0.092	-0.154		
msat	**0.194***	**0.191***	**0.289***	**0.075***		
Madagascar: Belo sur Mer						
COI	0.176	0.157	0.150	0.164	0.178	
msat	**0.065***	**0.241***	**0.088***	**0.218***	**0.267***	

5.1.2). No significant IBD effects were detected for either species [*L. mahsena,* R^2 = 0.071, *P* = 0.340][*L. harak,* R^2 = 0.434, *P* = 0.090]. BAPS analysis detected two genetically distinct groups within the *L. harak* dataset, with Mauritius separated from the other SWIO samples (Supplementary Figure 5.1.1); however, FCA analysis divided samples into four groups corresponding with sampling sites (Fig.5). For *L. mahsena*, BAPS analysis identified 3 groups with samples from the Seychelles, Mozambique and the rest of the SWIO each appearing to fall within distinct populations (Supplementary Figure 5.1.1) supported by FCA analysis (Fig. 5.1.5).

Values of relatedness based on r_{gq} fell within the 95% confidence limits expected for panmictic populations for all samples of *L. harak* (Supplementary Figure 5.1.2), however mean values greater than expected under panmixia was identified in *L. mahsena* (Supplementary Figure 5.1.3). Maximum-likelihood inferences of relatedness amongst pairs of individuals (Supplementary Figure 5.1.4 & 5.1.5) identified high levels of ambiguous kinship inference (i.e. individuals could not be unambiguously classified as related or unrelated for both *L. harak* (88.54%) and *L. mahsena* (79.56%) but did identify related dyads within all samples for both species except Madagascar, Belo sur Mer for *L. mahsena* .

5.1.3.4 Palaeodistribution modelling

Mean AUC values were 0.985 (S.D. = 0.003) and 0.980 (S.D. =0.016) for *L. harak* and *L. mahsena*, respectively, indicating very good model performance. Present day models for both *L. harak* and *L. mahsena* appear to accurately reflect the known contemporary ranges of both

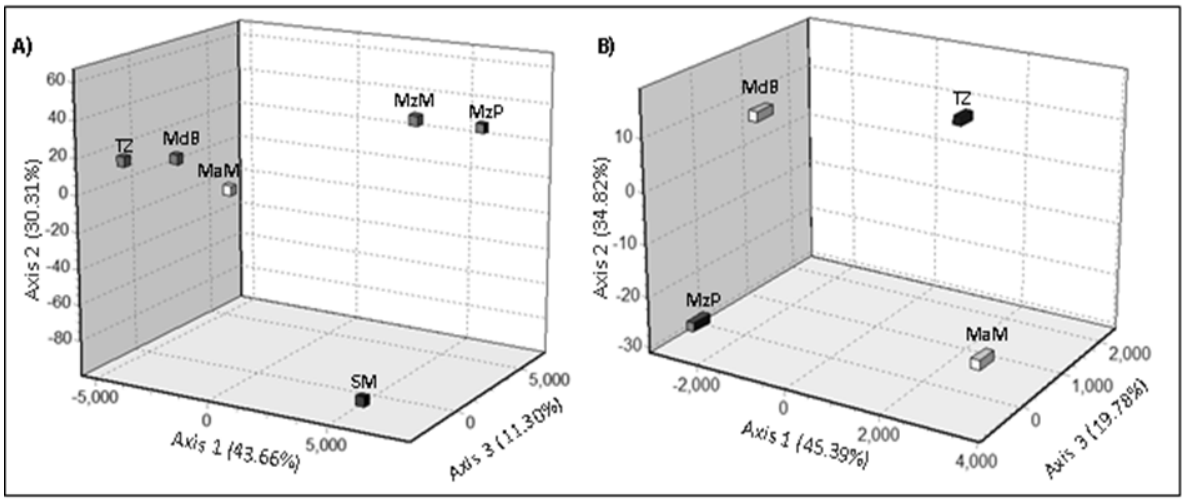


Figure 5.1.5: Factorial correspondence analysis showing multivariate relationships among microsatellite genotypes of SWIO samples of A) *Lethrinus mahsena* and B) *Lethrinus harak*

species, with suitable habitats aggregated in coastal environments across the Indian and western Pacific Oceans (Fig. 5.1.6). Despite an overall reduction in suitable habitat during the LGM, this appears to take the form of a geographically homogeneous contraction across both species ranges rather than obvious vicariant fragmentation of present day ranges (Fig. 5.1.6).

5.1.4 Discussion

5.1.4.1 Cryptic sympatric species

An important result was the detection of two highly divergent mtDNA clades (A and B) co-occurring within samples all described initially as *L. mahsena* from two sites in the Seychelles. Targeted PCR-RFLP analysis of all of our samples suggests clade B was confined to the Seychelles samples. Although both clade A and clade B exhibited high sequence similarity with *L. mahsena* voucher specimens from Mozambique and India, respectively, sequence divergence between the two clades (mean 4.4%) was in excess of even conservative thresholds of species -level divergence (Roux *et al.* 2016; Zemlak et al. 2009). Microsatellite analysis also revealed significant nuclear differentiation among members of the two clades, with bi-parentally restricted gene flow also apparent in results of assignment tests (high levels of assignment to clade). Clade B sequences were more similar to sequences of *L. atkinsoni*, a species described from west Pacific waters but not in the WIO. Sequence divergence between clade B and *L. atkinsoni* (mean = 2.2 %; K2P distance =0.021) is in the grey zone of the species divergence continuum (Roux *et al*. 2016), suggesting that clade B may represent a highly divergent *L. atkinsoni* clade lying well outside of the currently described species distribution, or that clade B represents an undescribed species most closely related to *L. atkinsoni*. Overall, the cyto-

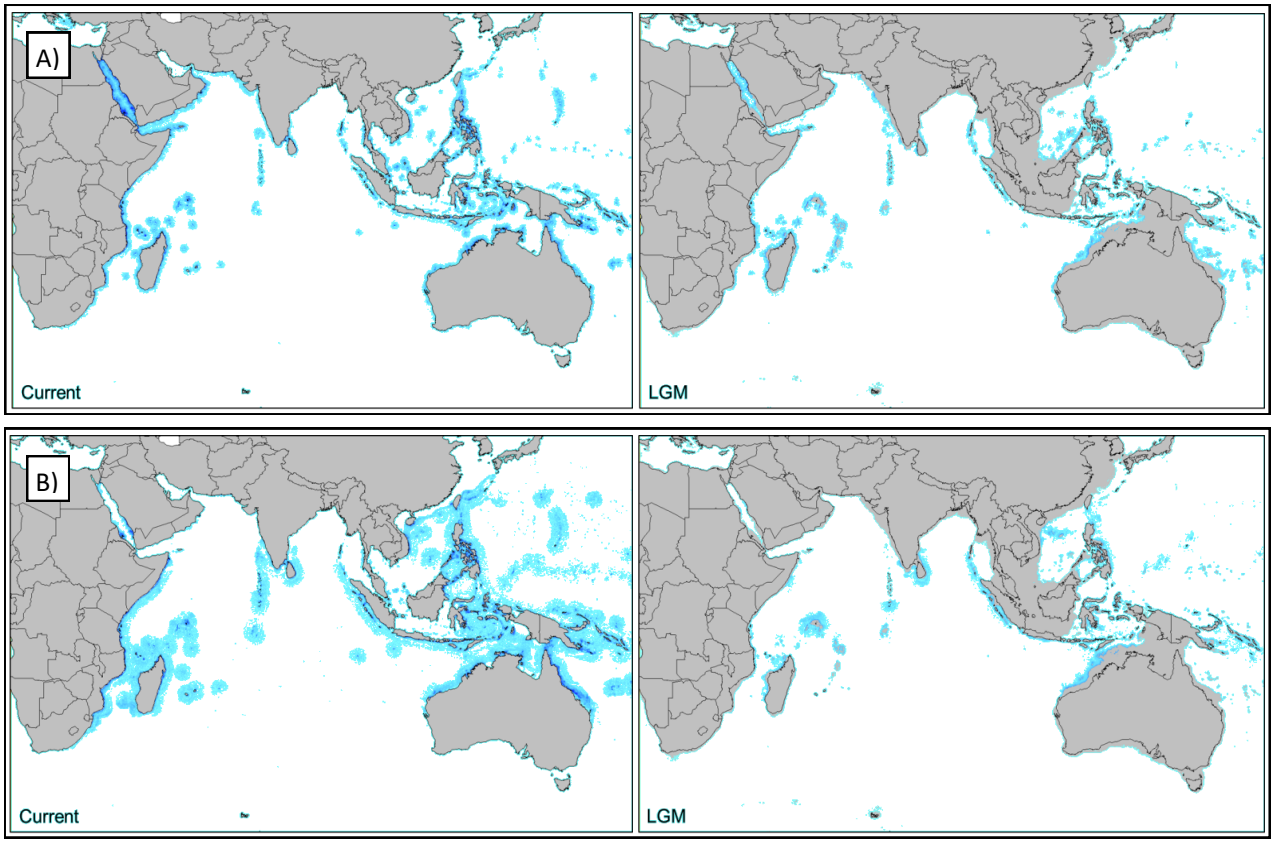


Figure 5.1.6: Results of species distribution models depicting predicted ranges of A) *Lethrinus mahsena* and B) *Lethrinus harak* during the present day and the Last Glacial Maximum (LGM, approximately 21KYA).

nuclear differentiation indicates the sympatric occurrence of a second species alongside *L. mahsena*, which is being indiscriminately harvested as *L. mahsena*. Furthermore, the data point to incongruence between current taxonomy and DNA barcodes. Specimens from India belonging to this potential cryptic species may have been misidentified as *L. mahsena* and used as voucher specimens. Further evidence of misidentification is apparent among sequences deposited as *L. mahsena* on GENBANK: those from Madagascar (JQ350088.1, JQ350089.1) have a greater similarity (99%) with *L. lentjan* voucher specimens from Iran, whereas supposed *L. mahsena* sequences from Japan (JF952782.1) exhibit greatest similarity (91%) *to L. rubriopercu-latus*. In general, the meristic and morphological features of emperor fish are conservative (Carpenter & Allen 1989) and distinct species are often difficult to identify easily by morphology alone (Sato 1971; Smith 1959). Failure to describe the full species diversity and/or misidentification of species may fundamentally compromise conservation strategies and tools. For example, the ENM performed here was based on reported occurrences of the species, however based on the genetic results such occurrence data may be inaccurate. By extension, the modelled habitat patches may actually reflect distinct species ranges. The cryptic biocomplexity resolved here, as well as other cases of identification of cryptic lethrinid species (Borsa *et al.* 2013), emphasises the need for genetic studies on a wider geographical scale to fully resolve

levels of intra-specific and inter-specific diversification, as well as for the development of low cost species identification and assignment techniques as resources for management such as the PCR-RFLP procedure developed here.

5.1.4.2 Population history of *Lethrinus harak* and *Lethrinus mahsena* across the SWIO

Hindcasting of ecological niche models (ENMs) suggested a marked reduction in suitable habitat for both *L. mahsena* and *L. harak* during the LGM that would be expected to have decreased their range-wide population sizes. Therefore, as for many other SWIO taxa including numerous reef associated fish (Craig et al.2007; Visram et al.2007) and crustaceans (Tolley et al.2005; Gopal et al.2006; Fratini et al.2010), it is likely that eustatic sea level fluctuations during the Pleistocene effected demographic changes in SWIO lethrinids. mtDNA phylogenies for both species conformed to classical 'star' shaped patterns (Slatkin & Hudson 1991) expected under population expansion models. Demographic tests also provided some support for population expansion events for *L. mahsena* but not *L. harak*, however, the resolution of these tests may have been limited by a combination of low numbers of informative sites and differing sample sizes. The lack of genetic breaks within the mtDNA phylogenies for both species is compatible with a lack of prolonged vicariance. Similarly shallow phylogenies have been reported in many other SWIO species. Therefore it would appear that for many species there has been little phylogeographic diversification within the SWIO (e.g. Hoareau *et al*.2012; Muths *et al*.2015) with the majority of cases of deep phylogeographic structure within the region due to colonisation of allochthonous lineages (e.g. Ragioneiri *et al.*2009; 2010; Silva *et al.*2010a).

5.1.4.3 Population structure of *L. harak* and *L mahsena* across the SWIO

Although mtDNA Φ_{ST} values were high for both species, pairwise comparisons between samples were largely non-significant. In contrast, highly significant nuclear (microsatellite) differentiation was reported among samples for both species. Such a pattern is compatible with the greater statistical power of the nuclear data set as estimated by POWIM. Similarly, Muths et al. (2011) detected significant fine scale genetic structuring in the WIO for *Myripristis berndti* while a previous mtDNA-based study indicated extensive connectivity (Craig et al. 2007). For both *L. harak* and *L. mahsena* almost all pairwise tests (F_{ST} and exact) were significant, including comparisons between nearby samples, with no clear geographical pattern or significant IBD (there was also no geographical pattern to the mtDNA ϕ_{ST} values). The structuring for both species was also significant upon conservative correction for null alleles (as discussed by Shaw et al. 2010), and was not driven by single locus effects (no loci identified as outliers). F_{ST} values were considerably higher for *L. mahsena* than for corresponding comparisons in *L. harak.* While POWSIM indicated a low type I error rate for both data sets in qualitative (allele fre-

quency) tests the smaller sample sizes for *L. mahsena* may have inflated F_{ST} values (Willing 2012). From a biological viewpoint the considerably greater abundance of *L. harak* may also point to smaller effective population sizes for *L. mahsena* which could predispose the species to greater genetic drift effects and hence larger F_{ST} even if migration rates are similar (e.g. Whiteley *et al.* 2004). Further insight into the drivers of genetic differentiation was provided by the kinship analyses, which examine how alleles are shared among individuals and provide an independent test of the hypothesis that structure as quantified by F_{ST}, which focuses on population allele frequencies, is a result of connectivity (Iacchei et al. 2013; Christie et al. 2010). Firstly, for *L. mahsena* mean kinship values for all samples exceeded those predicted within a non-structured system. Second, related individual pairs, and/or potentially related dyads (i.e. dyads that could not be unambiguously described as unrelated) were found in all samples for both species. While the large number of dyads that could not be unanimously classified to a single relationship category (i.e. unrelated or related) highlights the resolution threshold of the data, the results for both allele frequency and allele sharing analyses indicate that, even though sample sizes are small in some cases, the genetic heterogeneity cannot be dismissed as statistical noise but rather reflects some changes in the composition that may be a useful tool for better understanding recruitment dynamics and connectivity in these species (Selkoe et al. 2006; Knutsen et al. 2011).

The geographically patchy genetic structuring and kinship patterns could be generated by three non-mutually exclusive processes: population isolation, large variances in individual reproductive success (sweepstakes recruitment), limited mixing of larvae from genetically different sources (larval cohesion). Facets of the genetic structure exhibited some congruence with known biogeographic boundaries and patterns of population isolation inferred from other species. For *L. harak*, there was pronounced differentiation of the Mauritian population from the rest of the SWIO, supported by clustering and pairwise $F_{ST} > 0.03$ in all comparisons. Although not resolved in the clustering analysis, the Mauritius *L. mahsena* sample was also highly differentiated from all other samples (pairwise F_{ST} values > 0.04). Differentiation of Mascarene populations have been observed for several other reef-associated species (e.g. Muths *et al.*2011; 2015) and linked to a combination of large geographic distances from other coasts (>1000km), and a barrier effect due to the landmass of Madagascar. Genetic differentiation of the *L. mahsena* Seychelles sample was also supported by pairwise tests and clustering analysis. A combination of the South equatorial current and Equatorial counter current could serve to restrict connectivity between more western/southern locations as suggested by Muths *et al*., (2015). Such an isolating mechanism might also explain the seeming absence of clade B from other SWIO sites. The high fecundity and batch spawning behaviour of both species could be condu-

cive to sweepstakes effects while local seascape features may contribute to recruitment heterogeneity (McKeown *et al*. 2017b). The Mozambique Channel has a dominant seasonal anticyclonic cell at the northern entrance, as well as a succession of mesoscale anticyclonic and cyclonic eddies along the Mozambique coast (Schouten et al. 2003; Swart et al. 2010). Such features have been suggested to contribute to stochastic recruitment/genetic patchiness within the region in a number of species (Bourejea et al. 2007; Hoareau et al. 2013; Muths et al. 2015) and may contribute to the differentiation among samples from both species within the Mozambique Channel reported here. Even in the absence of genetically isolated source populations, larval cohesion (Selkoe et al. 2006) may enhance (Waples 2002), and be effectively indistinguishable from sweepstake effects (Turner et al. 2007). Berry et al. (2012) reported evidence for cohesion among larval and juvenile *L. nebulosus* but suggested that this was subsequently dissipated by cumulative dispersal of juveniles with no non-random cohesion detected by 8 years of age, by which time *L. nebulosus* are reproductively mature. It is possible that the cohesion reported here is similarly transient, in which case the patterns would reflect harvesting of individuals that may not have attained reproductive maturity, a practise that can severely compromise sustainability.

Disentangling the exact roles of recruitment heterogeneity and restricted gene flow will require more sensitive genetic assays and would benefit from analysis of age-segregated samples (Burford et al. 2011). However, the overall results indicate processes within the region that prevent genetic mixing, at least, on timescales of interest to fishery managers. These contrasts with the broad scale genetic homogeneity and subtle patchiness reported for *L. nebulosus* in Australian waters (Berry et al. 2012), and implicates regional seascape drivers within the SWIO. In these cases a spatial 'bet-hedging' approach is advised for marine resource management, including geographical dispersion of marine reserves if they are to be used (Larson & Julian 1999). The ability to predict population structure across taxa applies directly to the implementation of MPAs. *L. harak* and *L. mahsena* reported similar historical modelled habitat occupancy and phylogeographic patterns, while resolved genetic patchiness points to similar recurrent population processes in both species. Therefore, while we agree that making sweeping predictions from alleged model organisms can be dangerous (Bird et al. 2007), This similarity suggests that the application of information from such models to subsets of taxa may be highly useful in light of finite resources for conservation. Paradoxically, the identification of a cryptic sympatric species assemblage with little (if any) morphological differentiation, that is potentially endemic to the Seychelles, highlights the utility and necessity of genetic approaches to characterise baseline biodiversity in the region before such model-based methods are employed.

Published in: ICES (2018)
DOI:10.1093/icesjms/fsx245

CHAPTER 5.2

Genetic analysis reveals harvested *Lethrinus nebulosus* in the south-west Indian Ocean comprise two cryptic species

5.2.1 Introduction

Genetic studies have yielded many insights into marine biodiversity, with important findings including the detection of significant genetic population structuring (Shaw *et al.*, 1999; Knutsen *et al.*, 2011; McKeown *et al.,* 2015) and adaptation (Hemmer-Hansen *et al.,* 2007; McKeown *et al.,* 2017) in systems where high gene flow would be expected to prevent such differentiation. In addition to resolving intraspecific population structure, genetic surveys have revealed substantial cryptic species diversity in the marine realm (Thorpe *et al.,* 2000; Borsa, 2002). The recognition of such distinct components is now regarded as fundamental to sustainable management.

The blue or spangled emperor fish, *Lethrinus nebulosus* (Forsskål, 1775), occurs throughout the Indo-West Central Pacific (Fischer and Bianchi, 1984; Grandcourt *et al.*, 2006). Throughout its distribution, *L. nebulosus* is harvested by various artisanal, recreation and commercial fisheries (Fischer and Bianchi, 1984; Heileman et al., 2015). Although historically a minor component of some regional fisheries, the species is being exposed to increased fishing pressure as catches of other species decline (Mann, 2000). Many *L. nebulosus* fisheries are already heavily exploited, operating above optimal sustainable levels and showing evidence of both recruitment- and growth-overfishing (e.g. Grandcourt *et al.*, 2006). Fishing pressure, coupled with habitat loss and alteration, are proposed to have led to declines in the biomass and abundance (or the complete absence) of *L. nebulosus* from certain areas, as well as declines in yields and catch per unit effort across its distribution (Heileman et al., 2015). Overfishing can also affect ecological processes (Hutchinson and Reynolds, 2004) and drive genetic changes that can seriously compromise adaptability (Iles and Sinclair, 1982), resilience and productivity (Ryman *et al.,* 1995).

The initial objective of this research was to describe population genetic variation of *L. nebulosus* in the South West Indian Ocean (SWIO), a region described as a biodiversity hotspot (Ridgway and Sampayo, 2005) and an area wherein fishing pressure on *L. nebulosus* is predicted to increase. Specifically, we aimed to test the null hypothesis of high connectivity throughout the region and compare patterns of genetic variation to other geographical regions to investigate if there were any signals of genetic erosion. However nuclear and mitochondrial da-

ta revealed two highly divergent genetic groups indicating that, despite no *a priori* evidence hitherto described, *L. nebulosus* in the SWIO actually comprise two cryptic species that are being indiscriminately harvested.

5.2.2 Methods

5.2.2.1 Sample collection & DNA extraction

Sampling was co-ordinated by the SWIO Fisheries Project (van der Elst et al., 2009). Samples (fin/muscle tissue/gill raker clips fixed in 95% ethanol) were collected from individuals identified as *L. nebulosus* based on morphology between 2011 and 2012, either on-board commercial fishing vessels or at commercial or subsistence landing sites across the SWIO (Fig. 5.2.1). Total genomic DNA was extracted from samples using a combination of standard CTAB-chloroform/isoamylalcohol methods (Winnepenninckx *et al.*,1993) alongside Qiagen (Hilden, Germany) DNeasy Blood and Tissue and Sigma-Aldrich (St. Louis, Missouri) GenElute Mammalian Genomic DNA Miniprep kits.

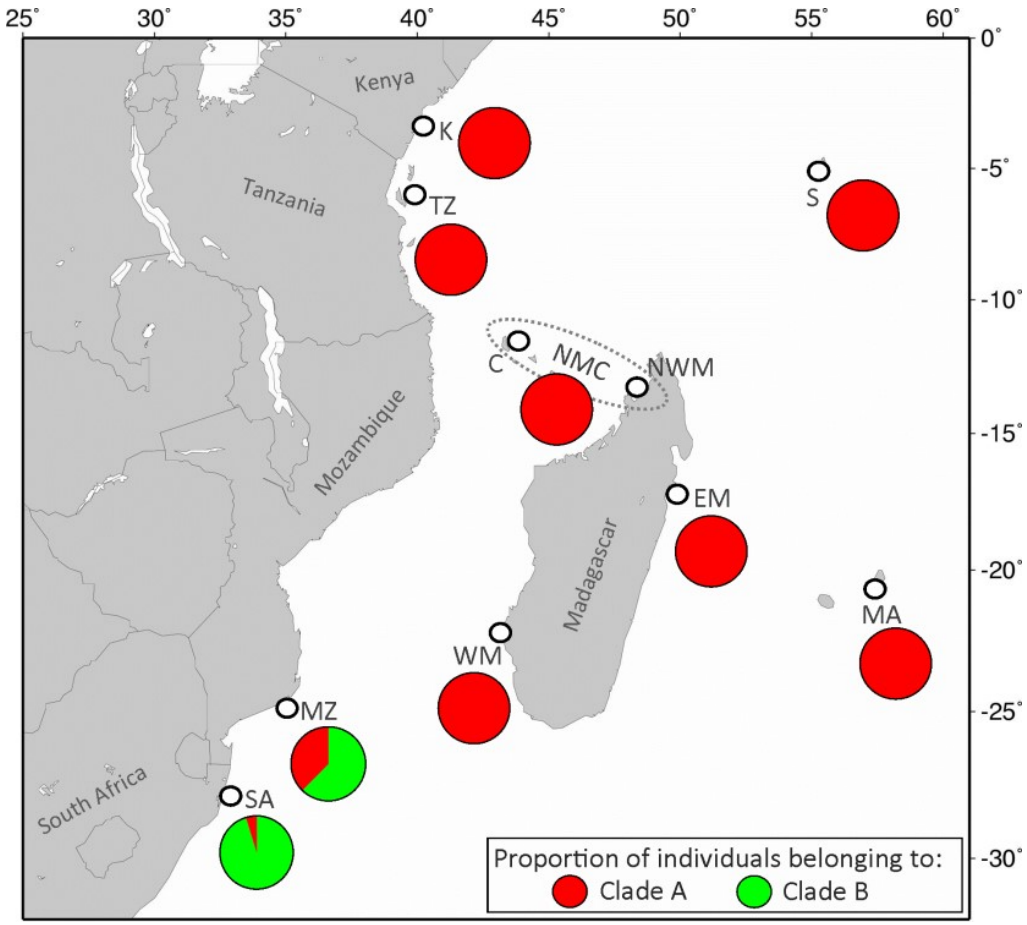


Figure 5.2.1: Sampling localities across the SWIO for L. nebulosus. Sampling locations are represented by circles and in some cases are midpoints of several geographically close sampling locations (see supplementary information for detailed sampling strategies). S=Seychelles, K=Kenya, TZ=Tanzania, C=Comoros, NWM =Northwest Madagascar, WM=West Madagascar, MZ= Mozambique, SA=South Africa, EM= East Madagascar, MA= Mauritius. Owing to small sample sizes C and NWM were grouped together and classed as NMC (Northern MozambiqueChannel). Pie charts represent the proportion of individuals in each sampling site that belonging to mtDNA clade A and B.

For genetic analyses, samples obtained from multiple localities within a country [e.g., Seychelles, (southern) Mozambique and South Africa] were pooled within each country. Samples from Madagascar were obtained from the east, north-west and south-west coasts. Owing to the longitudinal range of the island, those from the east and south-west coast were considered independently, whilst those from the north-west coast were combined with samples from the Comoros (owing to the small samples sizes from each of the individual localities) and designated as a Northern Mozambique Channel (NMC) regional sample (Fig. 5.2.1, Supplementary Table 5.2.2).

5.2.2.2 Control Region and Cytochrome oxidase I amplification & analysis

A fragment of the mtDNA Control Region (CR) was amplified by polymerase chain reaction (PCR) using species-specific primers (LCR1F 5'- CGGTCTTGTAAACCGGATGT-3' and LCR1R 5'-GTCATGGCCCTGAAATAGGA -3'). PCRs were performed in a total volume of 20 µL, containing 4 µL template DNA, 2 mM $MgCl_2$, 0.5 µM forward primer and 0.5 µM of reverse primer, 0.2 mM dNTP mix, 1x reaction buffer and 5U *Taq* polymerase (BIOTAQ). The PCR thermoprofile was: 180s at 95°C, followed by 45 cycles of 30s at 95°C, 45 seconds annealing at 54°C and 60s at 72° C, with a final extension period of 10 minute at 72°C. PCR products were then purified using EXOSAP-IT and sequenced in both directions on an Applied Biosystem 3500 platform using the respective PCR primers. Sequences were trimmed manually, and 421bp aligned across individuals using BIOEDIT (Hall, 1999). Additionally to facilitate a more comprehensive BLAST analysis a 443bp region of the barcoding COI gene was amplified in a subset of individuals representative of CR clades (described in results), using universal primers FishF1 and FishR1 (Ward *et al.*, 2005) and similar PCR condition as those for the CR, with annealing temperature reduced to 55°C.

Phylogenetic relationships among sequences were assessed using maximum likelihood (ML) and Bayesian inference (BI) trees, constructed in MEGA v7.0.21 (Tamura *et al.*, 2013) and MRBayes v3.2 (Ronquist and Huelsenbeck, 2003), respectively, using the optimum mutation model (HKY+G+I) inferred by JMODELtest (Posada, 2008). ML bootstrap (BS) values were calculated using 1000 bootstrap replicates. BI phylogenies were calculated assuming unknown model parameters, and run over 5,000,000 generations, sampling the Markov chain every 1000 generations, with the first 15% of trees discarded as burn-in. Median joining haplotype networks were calculated and constructed in NETWORK (Bandelt *et al.*, 1999).

Genetic variation was described using number of haplotypes (*H*) in addition to indices of haplotype (*h*) and nucleotide (π) diversity alongside their variances, all of which were calculated in Arlequin v3.5.2.2 (Excoffier and Lischer, 2010). Genetic differentiation among samples was

assessed using global and pairwise Φ_{ST} values, implemented in Arlequin with significance tested after 10,000 permutations.

5.2.2.3 Microsatellite amplification & analysis

Nuclear genetic variation was assessed at 14 microsatellite loci developed for *L. miniatus* (VanHerwerden *et al.*, 2000a,b), using protocols refined for *L. nebulosus* by Berry *et al.*, (2012). Fragments were sized, alleles scored and a genotype matrix produced using GeneMapper 5 (Applied Biosystems). Micro-Checker 2.2.3 (Van Oosterhout *et al.*, 2004) was used to test for signatures of technical artefacts. The number of alleles (NA), allelic richness (AR), the absolute number of private alleles (N_P), the percentage of polymorphic loci ($P_{\%}$) and observed (H_O) and expected (H_E) heterozygosities were calculated using GenAlex 6.502 (Peakall and Smouse, 2012). Deviations from Hardy-Weinburg (HW) expectations across loci and samples were assessed using exact tests in GENEPOP 4.5.1 (Rousset, 2008) and linkage disequilibrium was examined using exact tests in Arlequin.

Genetic structure was investigated without prior sample definition, using the model based Bayesian clustering programme; STRUCTURE V2.3.4 (Pritchard *et al.*, 2000). Data were explored using combinations of admixture/no admixture models with 500,000 MCMC repetitions and a burn-in of 100,000 iterations. Twenty replicates were conducted for each number of clusters (K), and the K value that best fit the data set was estimated using the log probability of data [PrX/K)] (Pritchard *et al*., 2000). Additionally the relationship among samples, based upon allele frequencies was visualised using factorial correspondence analysis (FCA) in GENETIX (Belkhir *et al.*, 2004).

To quantify genetic differentiation, pairwise F_{ST} values (Wright, 1951) were calculated in Arlequin, with significance determined after 10,000 permutations. Hierarchical analysis of genetic variance (AMOVA; Excoffier *et al.*, 1992) was performed in ARLEQUIN to quantify, and test significance using 1000 permutations of, variation among groups (F_{CT}) and among samples within groups (F_{SC}), with groups defined by mtDNA clade. To further investigate levels of nuclear differentiation between members of each mtDNA clade, 'leave one out' self-classification tests were conducted in GENECLASS 2 (Piry et al., 2004).

5.2.3 Results

5.2.3.1 mtDNA variability, phylogeny and BLAST results

An edited 421bp fragment of the CR was aligned across 292 individuals from nine samples within the SWIO, producing a total of 234 haplotypes. Phylogenetic reconstruction (Fig. 5.2.2, Fig. 5.2.3) revealed the presence of two highly divergent clades within the SWIO, separated by an average of 50.341 nucleotide differences (SE = 6.001). Average sequence divergence be-

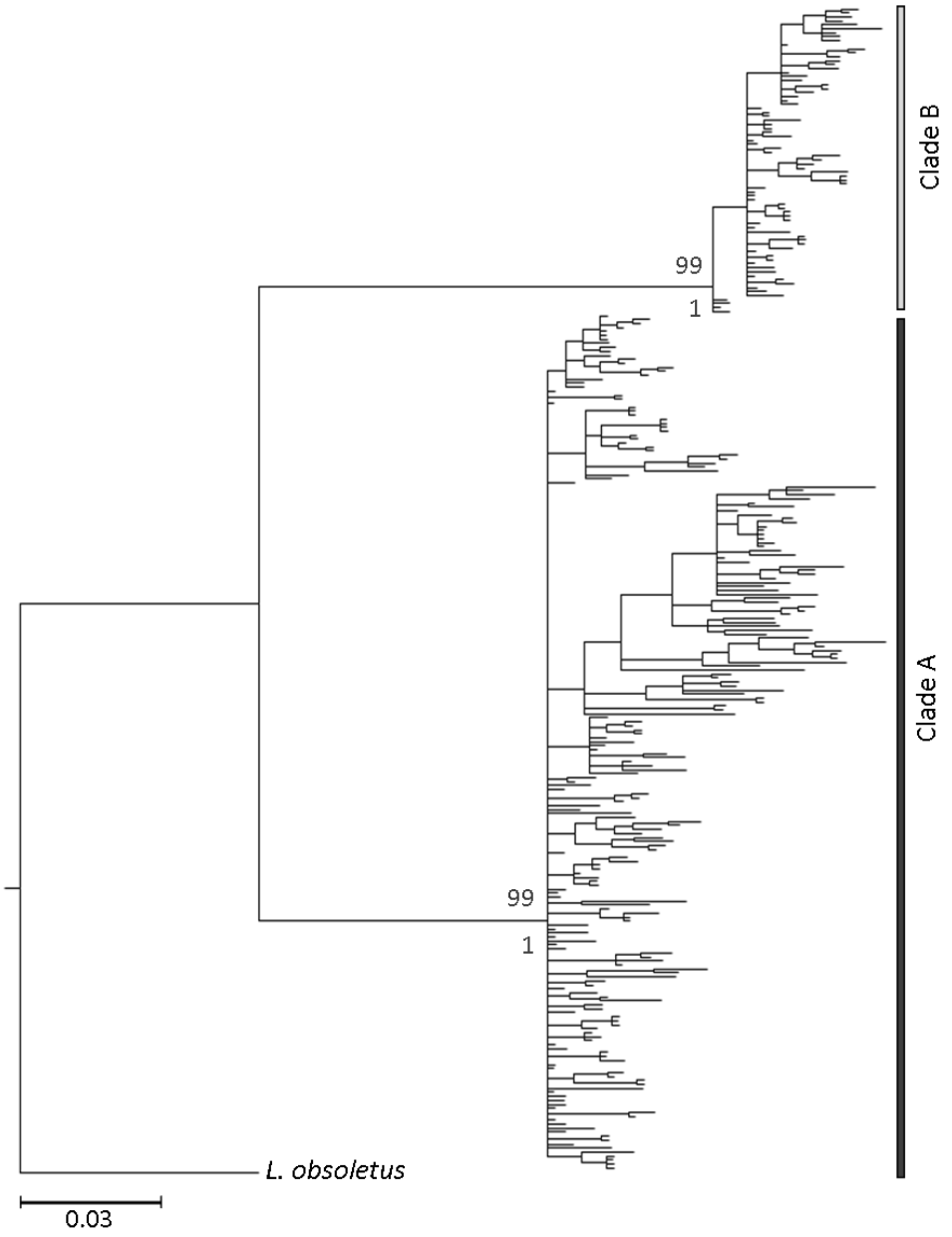


Figure 5.2.2: Phylogenetic relationships within SWIO *Lethrinus nebulosus* clades A and B, using 421bp CR mtDNA. Statistical support for nodes is given for both ML analyses (bootstrap support) above branches and Bayesian analyses (posterior probabilities) below branches. *Lethrinus obsoletus* is used as an outgroup.

tween Clade A and Clade B was 11.96%. Pairwise comparisons between samples belonging to each clade revealed large and significant Φ_{ST} values in all comparisons (Φ_{ST} = 0.822–0.889; Table 5.2.2). Overall haplotype diversity among clade A individuals was 0.998 (±0.001) and clade B was 0.991 (±0.004). Haplotype diversity for each clade was also high within each sample (Table 5.2.1).

Clade A was found across all SWIO samples, whilst clade B was found exclusively in Mozambique and South Africa where it co-occurred with and was more abundant than Clade A. Using BLAST, Clade A CR haplotypes reported an 89% sequence similarity to *L. nebulosus* from New Caledonia (EU983053, EU983038, EU983064, EU983049, EU983040, EU983084, EU983075, EU983045, EU983030, EU983024, EU983011), whereas Clade B bore an 84% sequence similar-

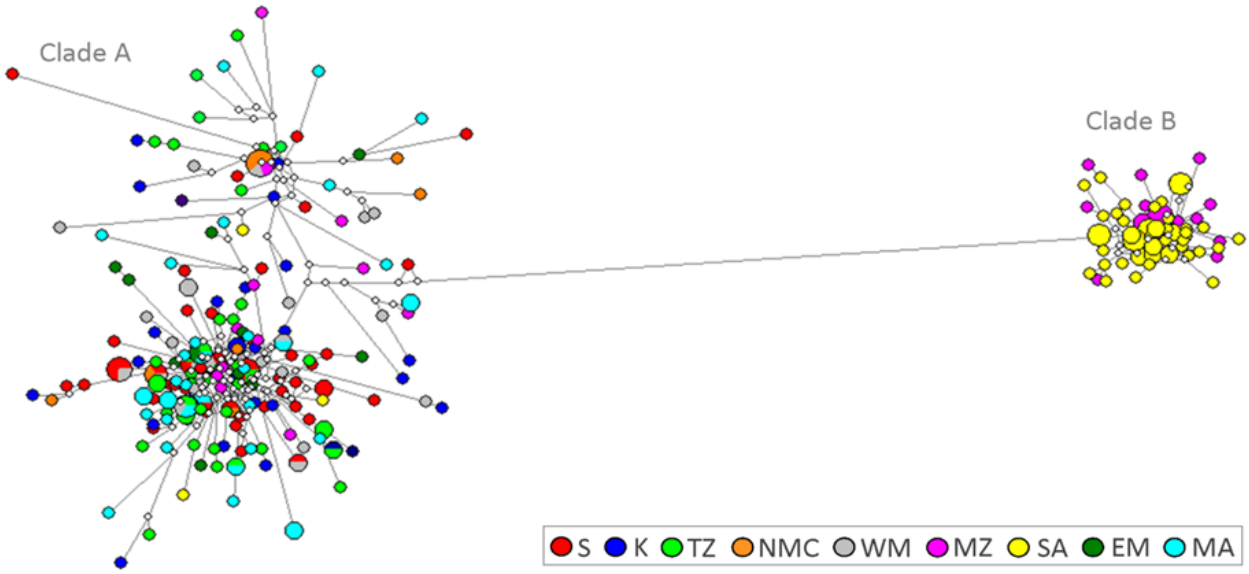


Figure 5.2.3: Median joining Haplotype network of *L. nebulosus* based on 421 bp CR mtDNA. Branch lengths are proportional to the number of differences. The node size is proportional to the haplotype frequency and node colour represents sampling location. Each small white circle represents a median vector. S=Seychelles, K=Kenya, TZ=Tanzania, NMC=Northern Mozambique channel, WM=West Madagascar,MZ=Mozambique, SA=South Africa, EM=East Madagascar, MA= Mauritius.

ity with *L. obsoletus* (AP009165) and *L. nebulosus* from New Caledonia and Bali (EU983030, EU983060, EU983027, EU983016, EU983090, EU983074, EU983040, EU983032, EU983020, EU983051).

The separation of clades was also apparent among the representative COI sequences which revealed an average inter-clade sequence divergence of 5.64% (Supplementary Table 5.2.3). Analysis with BLAST of representative clade A COI sequences reported 99% sequence similarity with voucher *L. nebulosus* sequences from Pomene in Mozambique (HQ561492, JF493754) and formed a monophyletic cluster alongside *L. nebulosus* sequences from across the Indian Ocean and west Pacific (Supplementary Figure 5.2.1). Similarly, BLAST analysis of clade B sequences reported 99% sequence similarity with voucher sequences of *L. nebulosus* from South Africa (DQ885022, DQ885021, DQ885020, JF493753). BLAST analysis revealed that both clades were more similar to *L. ornatus* than to each other (Clade A: sequence divergence = 5.66%; Clade B: sequence divergence = 3.16%).

5.2.3.2 Microsatellite variability and congruence with mtDNA

A total of 242 individuals, across eight samples (south-west Madagascar not genotyped), were genotyped at 14 microsatellite loci (Table 5.2.1). Microsatellite genotypes revealed clear partitioning of samples into two groups that aligned directly with clade membership. For example, the STRUCTURE analysis (Fig. 5.2.4a) identified the most probable model to be K = 2 (P = 1 and 0 for all other K values tested), wherein members of both clades were separated with high probability (Q values > 0.8). STRUCTURE analysis also did not reveal any individuals with inter-

Table 5.2.1: a) Mitochondrial genetic diversity levels and neutrality tests for *L. nebulosus* based on a) 421 bp of CR mtDNA: mtDNA sample size = *N*, number of haplotypes = *H*, number of private haplotypes = *pHap*, haplotype diversity = *h*, nucleotide diversity = π; and b) microsatellite genetic diversity across 14 loci in *L. nebulosus*: mean sample size per locus = *N* (± standard error; SE), mean number of alleles per locus = N_A (±SE), allelic richness = *AR*, the percentage polymorphic loci = *P%*, mean observed heterozygosity per locus = H_O (±SE),mean expected heterozygosity per locus = H_E (±SE), probability of deviation from Hardy Weinberg expectations = *P* and mean F_{IS} per locus = *F* (±SE). Statistically significant estimates ($P < 0.05$) are highlighted in bold. Instances where tests could not be performed are indicated as such (-). Standard deviation is given in parenthesis.

	a) mtDNA CR					a) msats					
	N	H	h	π	*N*	N_A	*AR*	*P%*	H_O	H_E	*P*	*F*
Seychelles	49	44	0.994 (0.006)	0.0370(0.0187)	36.143 ±0.206	11.500 ±2.062	3.637	100.00	0.699 ±0.063	0.697 ±0.064	0.258	0.000 ±0.035
Kenya	29	28	0.998 (0.010)	0.0498(0.0252)	28.214 ± 0.459	10.143 ±1.760	3.514	100.00	0.681 ±0.060	0.682 ±0.064	**<0.001**	-0.020 ±0.034
Tanzania	41	37	0.995 (0.007)	0.0399(0.0202)	28.429 ±0.644	9.929 ±1.746	3.545	100.00	0.684 ±0.051	0.696 ±0.058	**<0.001**	-0.007 ±0.042
NMC	10	7	0.867 (0.107)	0.0578(0.0315)	10.071 ±0.339	6.500 ±0.930	3.491	92.86	0.734 ±0.075	0.653 ±0.066	0.532	-0.130 ±0.032
W. Madagascar	22	21	0.996 (0.015)	0.0492(0.0253)	-	-	-	-	-	-	-	-
Mozambique (Clades A & B)	32	27	0.984 (0.014)	0.1094(0.0542)	31.929 ±0.071	12.714 ±2.484	-	100.00	0.670 ±0.071	0.714 ±0.070	**0.002**	0.056 ±0.033
Mozambique Clade A	12	12	1.000 (0.034)	0.0555(0.0297)	12.000 ±0.000	7.714 ±1.126	3.627	100.00	0.720 ±0.071	0.681 ±0.059	0.551	-0.052 ±0.075
Mozambique Clade B	20	15	0.958 (0.033)	0.0188(0.0103)	19.929 ±0.071	8.857 ±1.667	3.484	85.71	0.639 ±0.088	0.651 ±0.083	0.357	0.025 ±0.034
South Africa (Clades A & B)	60	45	0.989 (0.005)	0.0356(0.0179)	57.429 ±0.644	13.214 ±2.850	-	100.00	0.651 ±0.076	0.678 ±0.080	**<0.001**	0.029 ±0.024
South Africa Clade A	3	3	1.000 (0.272)	0.0544(0.0417)	3.000 ±0.000	3.214 ±0.408	3.214	85.71	0.524 ±0.125	0.536 ±0.074	0.859	0.093 ±0.155
South Africa Clade B	57	42	0.988 (0.006)	0.0199(0.0104)	54.429 ±0.644	11.786 ±2.553	3.515	100.00	0.658 ±0.077	0.671 ±0.080	**<0.001**	0.010 ±0.023
E. Madagascar	13	13	1.00 (0.030)	0.0505(0.0269)	11.929 ±0.071	7.857 ±1.378	3.599	100.00	0.660 ±0.074	0.658 ±0.068	0.311	-0.008 ±0.051
Mauritius	36	31	0.992 (0.008)	0.0471(0.0237)	29.286 ±0.244	10.571 ±2.027	3.634	92.86	0.683 ±0.071	0.693 ±0.071	**0.009**	0.007 ±0.037
Total	292	234	0.998 (0.001)	0.0757(0.0025)	233.429 ±2.104	19.643 ±3.679	-	100.00	0.677 ±0.063	0.728 ±0.068	**<0.001**	0.066 ±0.017

mediate assignment probabilities that might indicate hybrids or admixed genotypes. GENECLASS assignment tests also revealed 100% classification of individuals (165 for clade A and 77 for clade B) according to clade. This nuclear/mtDNA disequilibrium was also apparent in the FCA plots (Fig. 5.2.4b), which identified two distinct clusters of individuals corresponding to Clade A and B ancestry. Additionally, pairwise comparisons between samples partitioned according to mtDNA clade revealed high and significant F_{ST} values in all comparisons (F_{ST} = 0.084–0.115; Table 5.2.2). AMOVA of microsatellite genotypes partitioned according to mtDNA clade, revealed a significant component of the nuclear variation (8.84%, F_{CT} = 0.088; $P < 0.001$) present was attributed to differentiation among the two mtDNA clades with much lower, but still significant structuring within clades (0.78% F_{SC} = 0.009, $P < 0.001$).

3.3 Within clade structuring

As both mtDNA clades co-occurred among the Mozambique and South African samples, these individuals were partitioned by clade membership for subsequent tests of structure within each clade. For the microsatellite dataset no scoring errors attributed to large allele drop-out or stuttering were detected; however, several significant instances of the possible presence of

Table 5.2.2: Pairwise estimates of genetic differentiation between L. nebulosus samples based on (a) Pairwise θ_{ST} values based on 421 bp mtDNA (below the diagonal) and (b) Pairwise F_{ST} values based on 14 microsatellite loci (above the diagonal). S=Seychelles, K=Kenya, TZ=Tanzania, NMC=Northern Mozambique channel, WM=West Madagascar,MZ=Mozambique, SA=South Africa, EM=East Madagascar, MA= Mauritius.

		Clade A									Clade B	
		S	K	TZ	NMC	WM	MZ	SA	EM	MA	MZ	SA
	S	-	0.000	0.000	0.001		0.000	0.028	0.000	**0.013**	**0.085**	**0.090**
	K	**0.461**	-	0.000	0.008		0.015	0.005	0.000	**0.012**	**0.107**	**0.109**
	TZ	0.007	**0.437**	-	0.009		0.012	0.012	0.004	**0.023**	**0.100**	**0.103**
	NMC	**0.350**	**0.214**	**0.317**	-		0.019	0.042	0.011	**0.022**	**0.092**	**0.104**
Clade A	WM	0.008	**0.413**	0.012	**0.253**	-						
	MZ	**0.442**	0.029	**0.415**	**0.108**	**0.371**	-	0.038	0.013	**0.031**	**0.102**	**0.109**
	SA	**0.511**	0.010	**0.479**	0.150	**0.425**	0.071	-	0.023	0.017	**0.102**	**0.115**
	EM	**0.099**	**0.302**	**0.081**	**0.157**	**0.067**	**0.247**	**0.310**	-	0.009	**0.084**	**0.090**
	MA	**0.019**	**0.419**	0.003	**0.272**	0.010	**0.375**	**0.434**	**0.070**	-	**0.096**	**0.101**
	MZ	**0.852**	**0.822**	**0.846**	**0.843**	**0.839**	**0.837**	**0.889**	**0.847**	**0.827**	-	0.006
Clade B	SA	**0.868**	**0.853**	**0.865**	**0.870**	**0.865**	**0.867**	**0.885**	**0.871**	**0.853**	**0.319**	-

null alleles (11 of 140 loci by sample comparisons) and linkage disequilibrium (53 of 1007 loci by loci by sample comparisons) were detected. Fifteen of 140 locus by sample comparisons showed significant ($P < 0.05$) deviations from HW expectations. However significant deviations from HW equilibrium and linkage disequilibrium were distributed randomly across loci and samples and so all loci were included in subsequent analyses based on *a priori* sample definition.

Both clades exhibited some degree of intra-clade structuring which was more apparent at mtDNA than nuclear loci. Within clade B, significant pairwise θ_{ST} values (Table 5.2.2) were observed between Mozambique and South Africa in the CR data-set ($\theta_{ST} = 0.319$), but corresponding microsatellite F_{ST} values were not significant ($F_{ST} = 0.006$). Within clade A, significant pairwise θ_{ST} values (Table 5.2.2) were observed between the majority of samples (excluding Seychelles-Tanzania, Seychelles-west Mozambique, Kenya-Mozambique, Kenya-South Africa, Tanzania-west Madagascar, Tanzania- Mauritius, west Madagascar-Mauritius, Mozambique-South Africa and South Africa-North Mozambique Channel). However, pairwise comparisons of microsatellite genotypes identified broad scale genetic homogeneity with the only significant F_{ST} values identified in comparisons between Mauritius and all SWIO samples, excluding east Madagascar and South Africa.

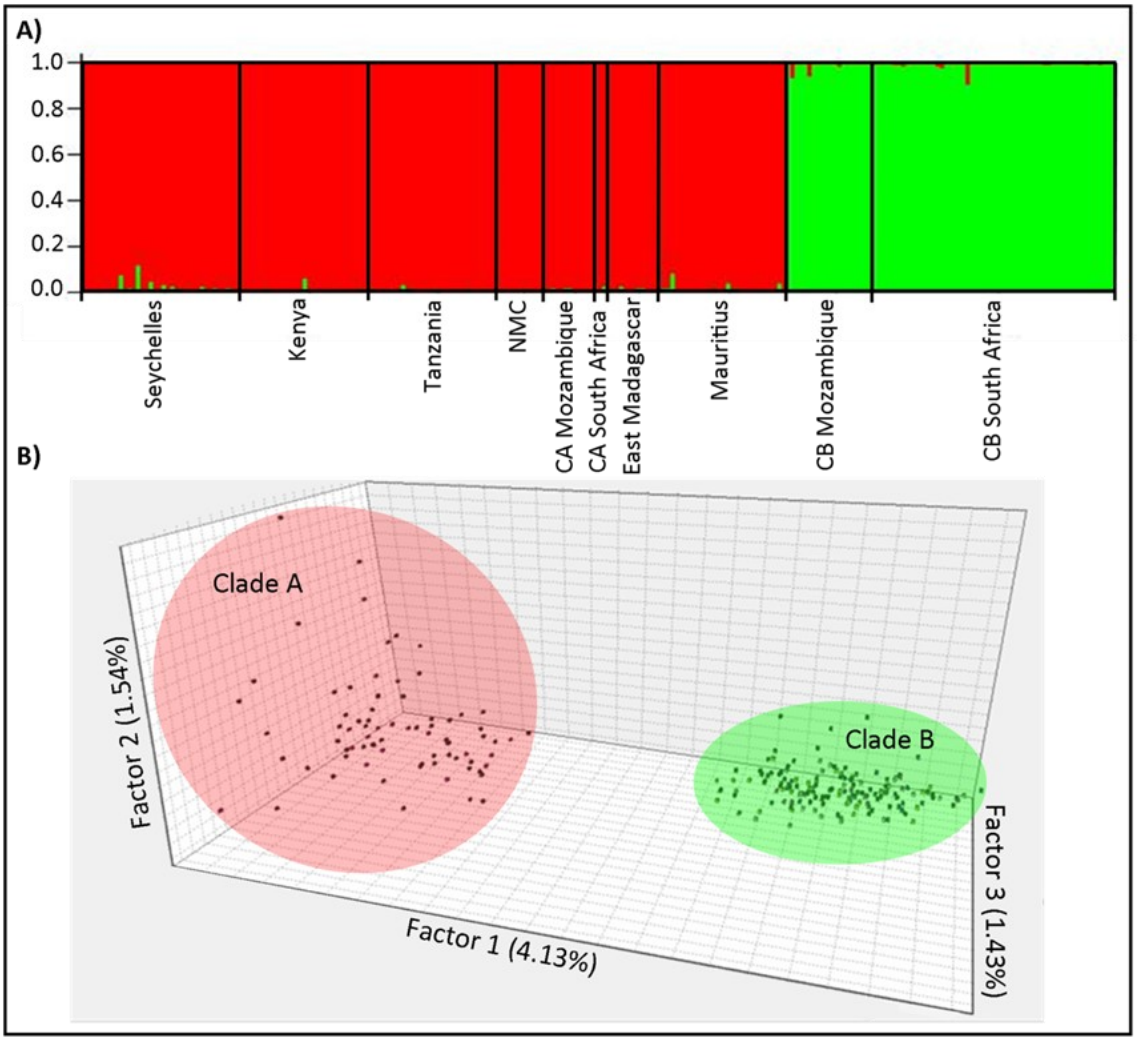


Figure 5.2.4: Estimates of genetic structure among 242 *Lethrinus nebulosus* individuals from eight regions in the south western Indian Ocean, based on genotypes from 14 polymorphic microsatellite loci. (a) STRUCTURE assignment of individuals to each of the two genetic groups (K=2) identified. Each vertical line represents an individual with shading corresponding to the probability of that individual (or the proportion of the individual's genotypic ancestry) belonging to Population A or Population B. (b) Plot of individuals along three factors identified in the Factorial Correspondence Analyses, identifying two clusters corresponding to groups A and B.

5.2.4 Discussion

The salient feature of this study was the concordant mtDNA and nuclear partitioning of samples into two distinct genetic groups. MtDNA revealed two reciprocally monophyletic clades. Average sequence divergence between clades (11.96%,) far exceeded that within clades A (2.95%) and B (1.09%), with this ratio well in excess of even conservative barcode gaps (Hebert *et al.,* 2004; Lefebure *et al.,* 2006). MtDNA data therefore indicate the occurrence of two highly divergent clades likely representing cryptic species. A common criticism of mtDNA-based taxonomy is that it represents a single, maternally inherited, locus (Collins and Cruickshank, 2013). However, analysis of nuclear microsatellites revealed robust partitioning of members of both clades with no evidence of hybrids, supporting a high level of bi-parental reproductive isolation.

There was a clear geographical pattern with clade A found throughout all SWIO samples whilst Clade B was restricted to locations at the south-western perimeter of the SWIO, where it co-occurred with clade A but at a higher frequency (at least in the samples collected here). This pattern was also concordant with geo-referenced samples identified through BLAST, where Clade B sequences clustered with *L. nebulosus* sequences from the western coast of South Africa, whereas Clade A sequences showed greater similarity with samples from across the Indian Ocean and west Pacific. The co-occurrence of clades A and B in southern Mozambique and north-east South Africa indicate a degree of spatial overlap in these clades. Bayesian clustering analysis indicated that this reflected mechanical mixing that was not accompanied by introgressive hybridisation. This clearly highlights the potential for indiscriminate harvesting of both species at local levels. Similarly, another genetic study of SWIO lethrinids (Healey *et al.*, in review) identified the co-occurrence of a cryptic species among samples presumed to be entirely composed of, and being harvested as *L. mahsena*. Such species misidentification may compromise sustainable management by leading to overestimation of stock abundances, or inappropriate management measures based on incorrect biological parameters such as age at maturity (e.g. Griffiths and Heemstra, 1995). On a wider scale, it also contributes to a fundamental underrepresentation of species richness and can cause inaccuracies in our understanding of biological, ecological and evolutionary processes (Garcia-Vazquez *et al.,* 2012).

These data add to the number of genetic studies that have identified cryptic species among lethrinids (Borsa *et al.*, 2013; Healey *et al.* in review) and it has been suggested that the conservative phenotypes of the family may have contributed to such undetected diversity. In this regard *L. nebulosus* is interesting as it exhibits a highly distinct phenotype relative to other lethrinids (Carpenter and Allen, 1989), hence it is frequently referred to as the “spangled” emperor. Against the background of morphological similarity, the finding of two species with a similar, highly distinct phenotype, that are not each other’s closest phylogenetic relative points to a mosaic of adaptive and plastic convergent effects.

For both species, microsatellites revealed limited evidence of structuring. Such broad-scale homogeneity is compatible with microsatellite patterns reported by Berry *et al.* (2012) for *L. nebulosus* in Australian waters. There was however some evidence of nuclear differentiation of Mauritian samples. Similar signals of isolation of Mauritian populations have been reported in other SWIO taxa, and have been attributed to a combination of large geographic distances from other coasts (>1000km), and a barrier effect due to the landmass of Madagascar (e.g. Muths *et al.*, 2011, 2015).

In contrast to the nuclear diversity, mtDNA revealed a pattern of chaotic genetic patchiness, common among marine species (e.g. Selkoe *et al.*, 2006). This difference between marker types may reflect the greater susceptibility of mtDNA to genetic drift. In line with this, genetic patchiness has been variously attributed to episodic drift effects due to large variances in individual reproductive success (sweepstake recruitment) and/or larval cohesion. Sweepstakes recruitment has been reported for a number of highly fecund taxa and may generate genetic differentiation despite gene flow when recruitment is variable (McKeown et al., 2017). Even in the absence of genetically isolated source populations, as might be the case here, larval cohesion (Selkoe *et al.,* 2006) may enhance, and be effectively indistinguishable from sweepstake effects (Turner *et al.*, 2007). Berry *et al.* (2012), through analysis of age segregated samples, empirically demonstrated the occurrence of larval cohesion in *L. nebulosus* and described how such cohesion, and associated genetic differentiation, were eroded by dispersal of sexually mature adults. As the samples analysed here are presumed to comprise multiple adults the detection of such differentiation may indicate that samples are being harvested prior to sexual maturity, a practise likely to compromise sustainability. Similarly, patchiness among mixed adult samples must be considered a conservative reflection of any underlying recruitment variability (McKeown *et al.*, 2017).

This study has important implications for management and conservation of lethrinids in the SWIO. Firstly, the resolution of two cryptic species within *L. nebulosus* highlights the inaccuracy of baseline biodiversity data for this group/region. In addition, the co-occurrence of both species within the southern samples suggests that both species are likely being indiscriminately harvested in this region. Such misidentification within fishery landings or stock assessments can severely compromise stock sustainability (Taylor *et al*., 2012; McKeown *et al.*, 2015). Finally, stochastic recruitment suggested to underpin the observed genetic patchiness may decrease resilience of local stocks to fishing and increase unpredictability in recovery (Kuparinen *et al*., 2014), and will necessitate a tailoring of the spatial scale of management (spatial bet hedging) according to biological and physical drivers of such recruitment variability. In addition, the detection of such patchiness may indicate that current fishing practises need to be amended to preclude harvesting of individuals that have not reached maturity.

CHAPTER 6

Drivers of diversification in the *Dentex dentex* species complex

CHAPTER 6

Drivers of eco-evolutionary diversification amongst west African Sparidae (*Dentex dentex*, *D. barnardi* and *Cheimerius nufar*)

6.1 Introduction

Disentangling the mechanisms and drivers of diversification and speciation can be challenging in the marine environment, with the high dispersal potential of many marine species (Craig *et al.* 2007), combined with the scarcity of discernible barriers to gene flow (Palumbi 1994) resulting in the assumption that marine populations are demographically open and genetically homogeneous throughout their geographical distributions (Cowen *et al.* 2000; Féral 2002; Craig *et al.* 2007). However, there are numerous well recognised drivers of genetic divergence in the marine realm (reviewed in: Palumbi 1994), including environmental, ecological and life history factors that serve to promote reproductive isolation between developing species.

The contemporary oceanographic features and the geological history of a marine region can impact the spatial distribution, dispersal pathways and ultimately the diversification and speciation of the taxa inhabiting it (Perez-Losada *et al.* 2002; Lessios *et al.* 2003; Teske *et al.* 2008; Henriques *et al.* 2012). The eastern Atlantic Ocean harbours a complex circulatory system with a broad band of equatorial tropical water bound to the north (Canary Current Upwelling System-CCUS) and south (Benguela Upwelling System-BUS) by cool upwelling currents, and as such the region is characterised by complex oceanographic systems and boundaries exhibiting steep temperature, nutrient and salinity gradients (Longhurst 1962; Hutchings *et al.* 2009; El-Geiziry & Bryden 2010). Furthermore, the interaction of these complex systems with historical environment and climatic changes are thought to have driven great temporal fluctuations in the marine environment across the eastern Atlantic (Diester-Haass *et al.* 1990; Krammer *et al.* 2006). All of which have been shown to have promoted vicariance and driven diversification across a range of eastern Atlantic taxa (Henriques *et al.* 2012; 2014; 2016; Gwilliam *et al.* 2018; Reid *et al.* 2016).

As a major component of coastal fish communities across the Mediterranean-north east Atlantic, west Africa and South Africa, the Sparidae represent an ideal group of species to investigate the roles of historical and contemporary vicariance in shaping the distribution and timing of evolutionary diversification. Here we use the *Dentex dentex* species complex as a model system. Based on morphology and molecular phylogenetics (Chiba *et al.* 2009; Sanitini *et al.* 2014; Carpenter & Johnson 2001; Orrell *et al.* 2002; Potts pers. comm) this species complex is

thought to consist of: 1) *Dentex dentex*, distributed throughout the Mediterranean and NE Atlantic as far north as the British Isles and south as far as Senegal (Bauchot & Hureau 1986; Morlaes-Nin & Moranta 1997; Bat *et al.* 2005), 2) *Dentex barnardi*, the distribution of which is thought to extend through the southern tropics, from Gabon to southern Angola (Bauchot & Hureau 1986; Bianchi *et al.* 1992; a;b); and 3) *Cheimerius nufar*, distributed across the warm temperate and subtropical coasts of South Africa (eastwards from Mossel bay) and into the tropical western Indian Ocean and Red Sea (Bauchot & Smith 1984; Fischer & Bianchi 1984; Heemstra & Heemstra 2004). Additionally, the molecular genetic study of Viret *et al.* (2018) identified an unknown Sparid linage in Portuguese samples (previously characterised as a divergent *D. dentex* lineage in Bargelloni *et al.* 2003), which clustered within the *D. dentex* species complex, forming a closely related and reciprocally monophyletic group with *C. nufar*. Although Virets *et al.* (2018) study lacked samples representative of *D. barnardi*, they suggest that in contrast to previous expectations (Chiba *et al.* 2009; Sanitini *et al.* 2014; Carpenter & Johnson 2001; Potts Pers comm), this unknown Sparid species constitutes the north-east Atlantic representative of this species complex, rather than *D. dentex*. *D. dentex, D. barnardi and C. nufar* all represent commercially attractive fisheries resources (Rueda & Martinez 2001; Duarte *et al.* 2005) and owing to their large size and high economic value represent a major component of commercial, recreational and artisanal fisheries across their distribution. However, the longevity (>30 years in *D. dentex*, >13 years in *D. barnardi* and >22 years in *C. nufar*), and protracted growth and maturation of these species render them vulnerable to overexploitation, as such the IUCN list *D. dentex* as "vulnerable" (Carpenter & Russell 2014) in their red list of threatened species.

Despite their commercial importance and vulnerability to overexploitation, *D. dentex, D, barnardi* and *C. nufar* remain relatively uncharacterised in regard to their evolutionary diversification and the spatial distribution of genetic variation, as exemplified by the resolution of cryptic diversity within this complex (Viret *et al.* 2018). Furthermore, and to the best of our knowledge there has been no molecular genetic investigation of the phylogenetic positioning or the evolutionary history of *D. barnardi*. As such this study initially set out to confirm the species status of *D. barnardi* and to provide molecular genetic corroboration of its assumed position (Carpenter & Johnson 2001) as a sister taxon to *C. nufar* within the Sparidae phylogeny. Secondly, through the incorporation of novel samples (i.e. west African *D. barnardi*) we aimed to assess the phylogenetic positioning and species status of the unknown Sparid linage identified by Viret *et al.* (2018) and assess its evolutionary inter-relationship with the *D. dentex* species complex. Finally, we aimed to explore the eco-evolutionary forces that have driven speciation across the Mediterranean, eastern Atlantic and Indian Oceans in these species,

and how they are related to the formation, location and geological history of known biogeographic barriers or dispersal pathways along the west African coast. Specifically, whether patterns of diversification conform more readily to those that would be expected based on vicariance or founder dispersal evolutionary scenarios.

6.2. Methods

6.2.1 Sample acquisition, DNA extraction, and PCR amplification

Samples were acquired through targeted fishing of adults along the Angolan and South African coasts, and supplemented in some places (Turkey) with individuals that were collected from local fish markets (Fig. 6.1). In total one sample (Turkey, Fethiye) was employed for *D.*

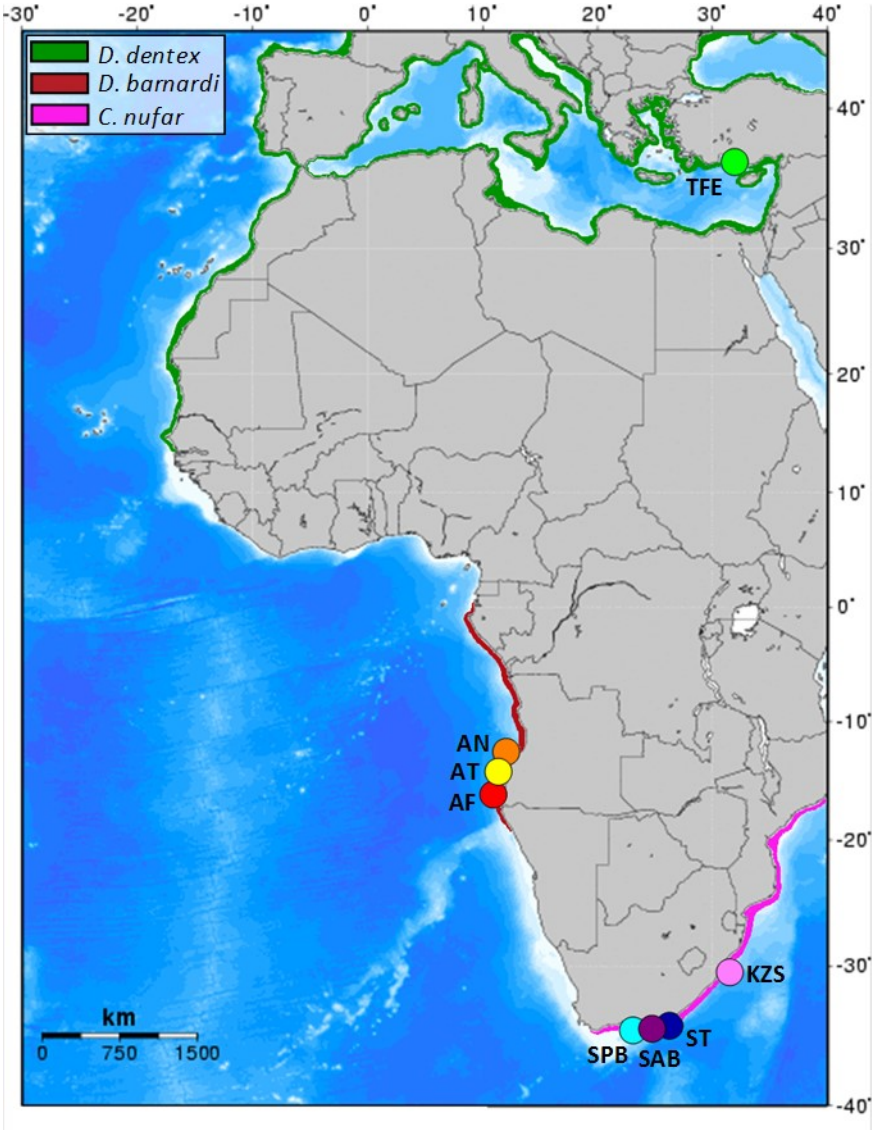


Figure 6.1: Sampling locations of Dentex species across South Africa, the west coast of Africa and the Mediterranean. TFE= sampling location of *D. dentex* in Fethiye, Turkey. AN, AT and AF represent Angolan sampling locations of *D. barnardi*, from Namibe, Tômbwa and Flamingo lodge respectively. SPB, SAB, ST and KZS represent South African sampling locations of *Cheimerius nufar* from Plettenburg Bay, Algoa Bay, Tsitsikamma national park and Scottburgh respectively. Coloured dots correspond to colours in Figs ? and coloured lines correspond to the recognised distribution of each species.

dentex, with three sampling locations across the Angolan coast chosen to represent *D. barnardi* and finally for *C. nufar* four locations were employed along the warm temperate eastern cape (Plettenburg Bay, Algoa Bay and Tsitsikamma national park) and tropical Indian Ocean coast (Scottburgh) of South Africa. Fin clips were taken from all individuals and stored in 95% ethanol, and DNA was extracted using a standard CTAB-chloroform/isoamyl alcohol method (Winnepennickx *et al.* 1993).

Two regions of the mitochondrial genome (COI and CR) were amplified by PCR. Firstly a fragment of cytochrome c oxidase subunit 1 (COI) gene was amplified using universal fish primers (FishF1:5'-TCAACCAACCACAAAGACATTGGCAC-3' and FishR1: 5'-TAGACTTCTGGGTGGCCAAA GAATCA-3' - Ward *et al.* 2005). Secondly primers were designed (DdenCR565F: 5'-CAGTAA GAACCGACCAACCT-3' & DdenCR16sR: 5'-TGGCAGAAAAGTCAGGACCA-3') from whole mitochondrial genome sequences of *Dentex dentex* and other closely related Sparids (Accession numbers: MG727892.1; MG653593.1; KT724963.1) to amplify the second hypervariable region of the Control Region (CR). PCRs were performed in a total volume of 20µl, containing 6µl of template DNA, 2 mmol/L $MgCl_2$, 0.5 µmol/L forward primer and 0.5 µmol/L of reverse primer, 0.2 mmol/L dNTP mix (20 µmol/L each dATP, dCTP, dGTP, dTTP), 1×reaction buffer [75 mmol/L Tris-HCl, mmol/L (NH4)2SO4], and *Taq* polymerase (BioTaq, 5 U/µl. The PCR thermoprofile for COI amplification was 3 minutes at 95°C, followed by 35 cycles of 95°C for 30 seconds, 1 minute annealing at 54°C and 1 minute extension at 72°C, with a final extension step of 3 minutes at 72°C. For CR amplification PCR conditions were: 3 minutes at 95°C, followed by 45 cycles of denaturing for 30 seconds at 95°C, 45 seconds annealing at 58°C and 1min extension at 72°C, followed by a final extension step at 72°C for 5 minutes. Additionally to explore patterns of nuclear partitioning for a subset of individuals a region of the alpha-amylase intron 1 locus of the nuclear genome was amplified, PCRs were performed using primers described in Hassan *et al.* (2002) (Am2B1F 3' CCTTCATCTTCCAGGAGGTAC'5; Am2B1R 3' TTCACCTCCCAGATCAATAAC'5), the volume of reagents and primers were identical to those used for the amplification of COI and CR. The PCR thermo-profile was: 5 minutes at 94°C, followed by 35 cycles of 94°C for 30 seconds, 45 seconds annealing at 55°C and 45 seconds extension at 72°C, followed by a final extension for 7 minutes at 72°C.

All PCR products were purified using SURECLEANPLUS and sequenced with forward and reverse primers using BIGDYE technology. Sequence quality was checked by eye and sequence identify confirmed by BLAST. Sequences were aligned using the CLUSTALW programme (Thompson *et al.* 1994), implemented in BIOEDIT (Hall 1999). Additionally COI and CR sequences from the same individual were combined to create a concatenated alignment of both mtDNA gene regions.

The nuclear alpha-amylase intron 1 region required more comprehensive quality control. As such, all bases were called manually and mixed peaks/ heterozygote sites were noted and given ambiguity codes. Of the 18 individuals amplified, 10 heterozygote sites were identified within 6 individuals. Nuclear alleles were then reconstructed using the PHASE algorithm (Stephens *et al.* 2001; Stephens & Donnelly 2003) implemented in DNASP (Librado & Rozas 2009), the resultant alignment was then condensed into alleles using DAMBE (Xia & Xie 2001), and this final alignment was used for further analyses.

6.2.2 Phylogenetic and phylogeographic analyses

Phylogenetic relationships across COI, CR and concatenated sequences of all three species (*D. dentex, D. barnardi* and *C. nufar*) were assessed using Maximum Likelihood (ML) and Bayesian Inference (BI) topologies, constructed in PHYML (Guindon *et al.* 2005) and MrBAYES (Ronquist & Huesenbeck 2003) respectively. Optimum substitution models for each mtDNA locus were inferred based on AIC values (Akaikine 1974) implemented in ModelTest, with the HKY+G model identified as the best fitting model for the COI dataset and HKY+G+I identified as optimal for the CR and concatenated datasets. ML bootstrap (BS) and aLRT support were calculated after 1000 bootstrap replicates. Bayesian inference phylogenies and posterior probabilities were run over 5,000,000 generations using 3 heated chains and one cold chain, with the first 15% of trees discarded as burn-in and the Markov chain sampled every 1000 generations. Convergence was considered reached when the standard deviation of split frequencies fell below 0.01. To further assess patterns of phylogenetic partitioning amongst and within species Median-Joining (MJ) haplotype networks were calculated and constructed in NETWORK (Bandelt *et al.* 1999). For the Alpha-amylase intron 1 dataset, MJ haplotype networks were constructed in NETWORK, using the PHASED alignment based on the most likely alleles identified in DAMBE.

Genetic variation and diversity for both mtDNA markers employed here (COI & CR) were examined using the numbers of haplotypes (H) alongside measures of haplotype (*h*) and nucleotide (π) diversity and their variances, calculated in Arlequin v3.5.2.2 (Excoffier & Lischer 2010). Genetic differentiation amongst all samples (within and across species) was assessed using global and pairwise Φ_{ST}, implemented in Arlequin v3.5.2.2 (Excoffier & Lischer 2010), with significance calculated after 10,000 permutations. COI and CR sequences were then partitioned based on their affinity to mtDNA lineage/species and genetic distance between species was assessed in MEGA v7.0.21 (Tamura *et al.* 2013), using the Kimura-2-Paramater+G model and the Tamura-3-Parameter+G+I models for COI and CR respectively, as these were the optimum models available for these analyses. To account for variability in model choice genetic dis-

tance was also assessed using p-distance calculated in MEGA v7.0.21. Finally the proportion of genetic variance distributed amongst species was estimated by analysis of molecular variance (AMOVA: Weir and Cockerham, 1984; Excoffier et al. 1992; Weir, 1996) implemented in Arlequin v3.5.2.2 (Excoffier & Lischer 2010), with significance tested after 1000 permutations.

6.2.3 Divergence dating

COI sequences generated here for *D. dentex, D. barnardi* and *C. nufar* were aligned with those employed in Viret *et al.* (2018) and Bargelloni *et al.* (2003), resulting in a 533bp long alignment from which a ML phylogeny was constructed. Estimated divergence time between all the major nodes of this phylogeny were calculated using the RelTime (Tamura *et al.* 2012; 2013) method, implemented in MEGA v7 (see chapter 3.1.2.2 for a more comprehensive discussion of this technique). RelTime was run using the HKY+G substitution model and calibrated using a fossil calibrations outlined in Santini *et al.* (2014).

Table 6.1: Measures of genetic diversity and neutrality across Dentex species for the COI and CR datasets. N= number of samples, H= number of haplotypes, *h*= haplotype diversity, π= nucleotide diversity, D=Tajima's D test for selective neutrality, Fs= Fu's Fs test of selective neutrality. Codes for sampling locations correspond to Fig 1. Standard deviations for haplotype and nucleotide diversities given in parentheses.

	N		H		*h*		π		D		Fs	
	COI	CR	COI	CR	COI	CR	COI	CR	COI	CR	COI	CR
TFE	16	16	3	11	0.492 (0.117)	0.950 (0.036)	0.00097 (0.00095)	0.01030 (0.0060)	-0.330	-0.928	-0.290	-3.780*
AN	18	18	7	17	0.634 (0.127)	0.994 (0.021)	0.00161 (0.00133)	0.0255 (0.0137)	-1.933*	-0.757	-4.322*	-7.422*
AT	14	13	5	11	0.593 (0.144)	0.974 (0.039)	0.00154 (0.00131)	0.0225 (0.0125)	-1.623*	-0.524	-2.100*	-2.373
AF	3	3	2	3	0.667 (0.314)	1.000 (0.272)	0.00246 (0.00252)	0.0185 (0.0149)	-	0.000	1.061	0.807
SPB	8	5	4	5	0.643 (0.184)	0.857 (0.137)	0.00263 (0.00203)	0.0224 (0.0135)	-1.175	-0.125	-0.519	1.419
SAB	10	9	4	9	0.533 (0.180)	1.000 (0.052)	0.00233 (0.00181)	0.0229 (0.0132)	-1.136	-0.348	-0.384	-3.064
ST	5	5	3	5	0.700 (0.218)	1.000 (0.127)	0.00147 (0.00147)	0.0177 (0.0117)	-0.973	-1.045	-0.829	-0.832
KZS	18	15	5	15	0.771 (0.063)	0.974 (0.029)	0.00205 (0.00157)	0.0251 (0.0135)	-0.134	0.315	-0.989	-3.939*
TOTAL	92	89	21	66	0.873 (0.020)	0.988 (0.005)	0.03743 (0.00524)	0.06601 (0.00352)				

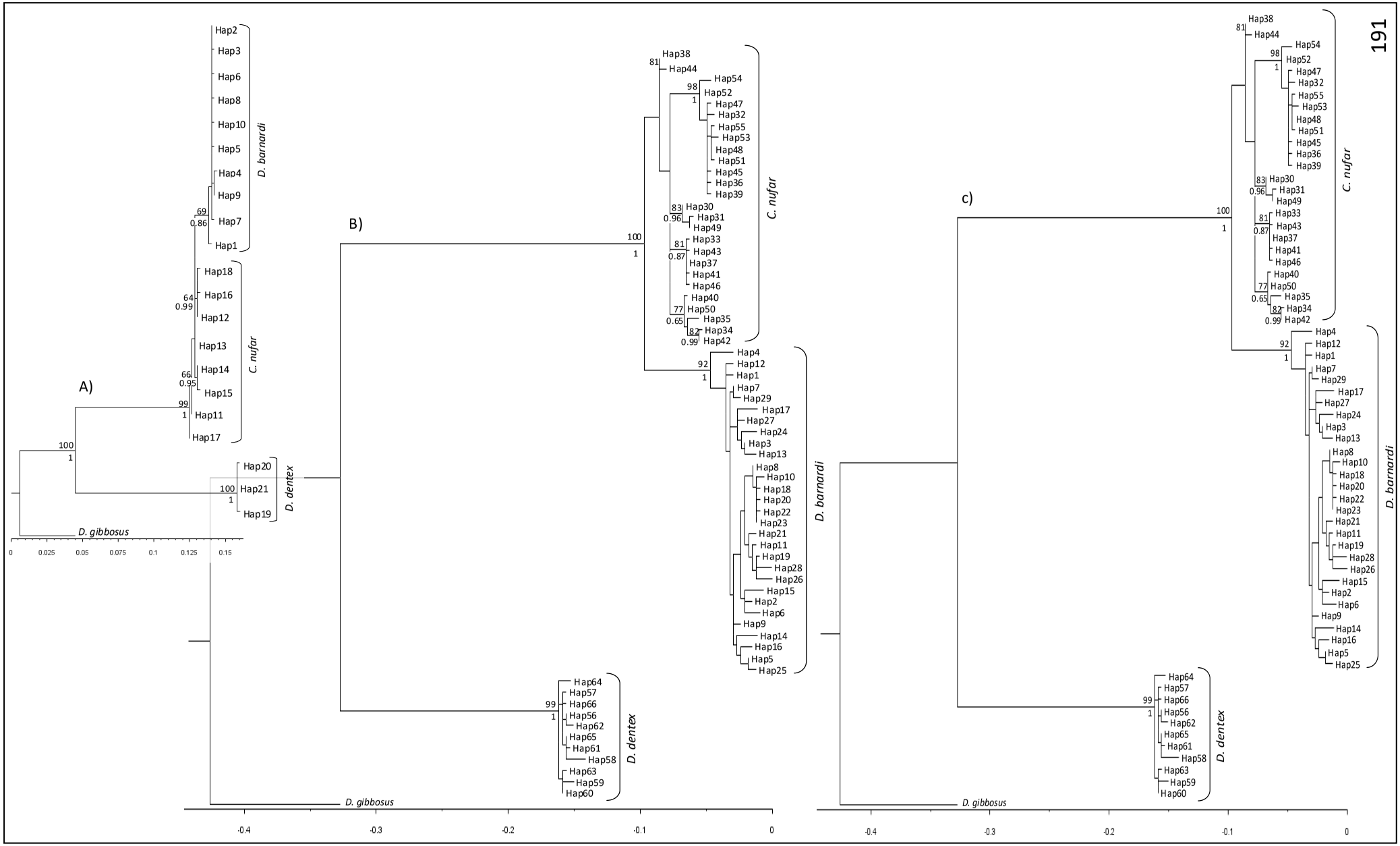


Figure 6.2: Maximum likelihood phylogenetic relationships amongst *D. dentex*, *C. nufar* and *D. barnardi* haplotypes based on A) 543bp of mtDNA COI, B) 398bp of mtDNA CR and C) 941bp of concatenated COI & CR. Statistical support for nodes is based on Bayesian analysis (BI) below branches and maximum likelihood bootstrap support (BS) above branches. Dentex *gibbosus* (accession number: MG653593.1) used as an outgroup.

6.3 Results

Across all 3 species included in this investigation (*D. dentex, D. barnardi* & *C. nufar*) a 543bp fragment of the COI gene and a 398bp fragment of the CR was aligned across 92 and 89 individuals respectively. When CR & COI sequences were combined the concatenated dataset consisted of a total of 941bp across 87 individuals. A 363bp (368bp including one 5bp indel that was present in a single individual from Scottburgh, KZN) fragment of the nuclear alpha-amylase intron 1 locus was aligned across a smaller subset of individuals (N=18).

6.3.1 Genetic diversity within and across species

Of the 92 individuals that made up the COI alignment, 21 haplotypes were resolved, within the CR data-set 66 haplotypes were resolved across 89 individuals (Table 6.1). Across all sampling locations *D. dentex* from Turkey exhibited the lowest nucleotide (COI=0.00097, CR=0.01030) and haplotype diversities (COI=0.492), excluding CR haplotype diversity, where *C. nufar* from Plettenburg Bay exhibited the lowest values (0.857). Within the COI dataset the highest haplotype diversity was observed within *C. nufar* sampled from Scottburgh (0.771) with the highest nucleotide diversity observed within *C. nufar* from Plettenburg bay (0.00263). Within the CR dataset the highest nucleotide diversity was reported for *D. barnardi* sampled from Namibe, and due to the high frequency of private haplotypes at CR haplotype diversities of 1.000 were observed in the Flamingo Lodge sample of *D. barnardi* and the Algoa Bay and Tsitsikamma National Park samples of *C. nufar*. When sequences were partitioned based upon morphological and phylogenetic species designations *Dentex dentex* exhibited the lowest haplotype (COI= 0.492, CR=0.950) and nucleotide diversities (COI=0.00097, CR=0.0103). Similarly *C. nufar* exhibited the highest haplotype (COI=0.713) and nucleotide (COI=0.00227, CR=0.0241) diversities, excluding haplotype diversity within the CR dataset where a marginally greater value was resolved for *D. barnardi* (0.988) than *C. nufar* (0.954).

6.3.2 Phylogenetic relationships amongst Dentex species

Reconstruction of the phylogenetic relationships of Dentex species through a combination of ML and BI analyses resolved similar topologies across markers. Three reciprocally monophyletic and highly supported lineages were resolved in the COI, CR and concatenated datasets (Fig. 6.2), corresponding to *D. dentex* (all samples from Turkey), *D. barnardi* (all samples from Angola) and *C. nufar* (all samples from South Africa). ML and BI (Fig. 6.2) topologies indicate a more ancient split between *D. dentex* and both west African and South African species, followed by a more recent (and shallow for the COI dataset) divergence between *D. barnardi* and *C. nufar*, the latter appearing to be the more basal (ancestral) of the two species. These patterns of divergence are further supported by MJ haplotype networks (Fig. 6.3) for COI, CR

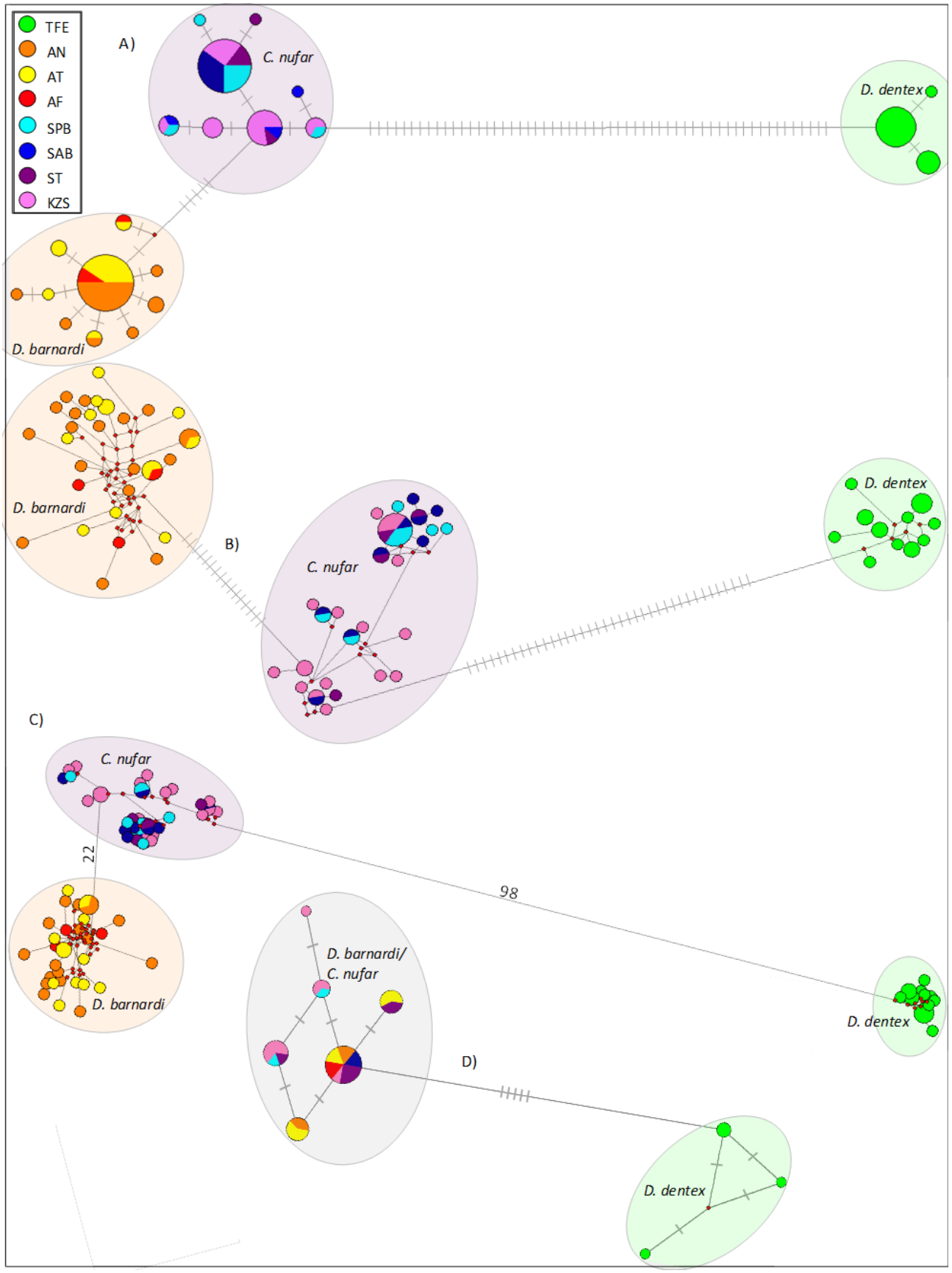


Figure 6.3: Median joining haplotype networks amongst *D. dentex*, *D. barnardi* and *C. nufar*, based on A) 543bp of mtDNA COI, B) 398bp of mtDNA CR, C) 941bp of concatenated COI & CR, and D) 363bp of the nuclear alpha-amylase intron 1 region. Branch lengths are proportional to the number of differences (each dash represents a single mutation) and node size is proportional to the haplotype frequency. Colours correspond to sampling location.

Table 6.2: Genetic distances between and within Dentex species calculated using A) 543bp of mtDNA COI, B) 398bp of mtDNA CR, C) 941bp of concatenated COI & CR. p distances given above diagonal, genetic distances calculated using the most appropriate substitution model (COI=Kimura-2-Parameter, CR = Tamura-3-parameter, concatenated = Kimura - 2 -parameter) given between species below the diagonal and within species on the diagonal.

A)	D. dentex	D. barnardi	C. nufar
D. dentex	**0.001 (0.001)**	0.113 (0.011)	0.109 (0.011)
D. barnardi	0.141 (0.018)	**0.002 (0.001)**	0.013 (0.005)
C. nufar	0.134 (0.018)	0.014 (0.004)	**0.002 (0.001)**
B)	**D. dentex**	**D. barnardi**	**C. nufar**
D. dentex	**0.010 (0.003)**	0.119 (0.016)	0.121 (0.014)
D. barnardi	0.143 (0.022)	**0.023 (0.005)**	0.067 (0.011)
C. nufar	0.146 (0.023)	0.075 (0.015)	**0.024 (0.004)**
C)	**D. dentex**	**D. barnardi**	**C. nufar**
D. dentex	**0.005 (0.001)**	0.115 (0.010)	0.114 (0.011)
D. barnardi	0.141 (0.015)	**0.010 (0.002)**	0.036 (0.005)
C. nufar	0.139 (0.014)	0.0138 (0.006)	**0.011 (0.002)**

and the concatenated datasets, where *D. dentex* is separated from *C. nufar* by 56, 38 and 98 mutations respectively, *C. nufar* is in turn separated from *D. barnardi* by 5 mutations at COI, 16 in CR and 22 in the concatenated COI/CR dataset.

Results of AMOVA (Table 6.3) and pairwise tests of genetic distance (Table 6.2) confirm that levels of divergence are far greater between than within species. AMOVA estimates found level of variability between species to be substantially greater than variability between populations within species and within populations.

No significant values of genetic differentiation (θ_{ST} –Table 6.4) were observed between sampling sites for any intra-specific comparison. When inspecting the concatenated *C. nufar* phylogeny (Fig. 6.4A) and MJ haplotype network (Fig. 6.4B), 5 clades (classified here as clades A-E) were resolved, all of which exhibited moderate to high statistical support (BS = 62-97, BI = 0.72-1.00). However, there appeared to be no clear geographical basis to the structuring of these clades.

Table 6.3: Percentage of COI, CR and concatenated mtDNA sequence variation explained by among -species, within-species and within-population variation as assessed by Analysis of Molecular Variance (AMOVA). Statistically significant results (p<0.05) in bold.

	COI	CR	concatenated
Among species	**96.8**	**77.74**	**86.50**
Within species	0.19	0.63	0.45
Within populations	**3.01**	**21.63**	**13.05**

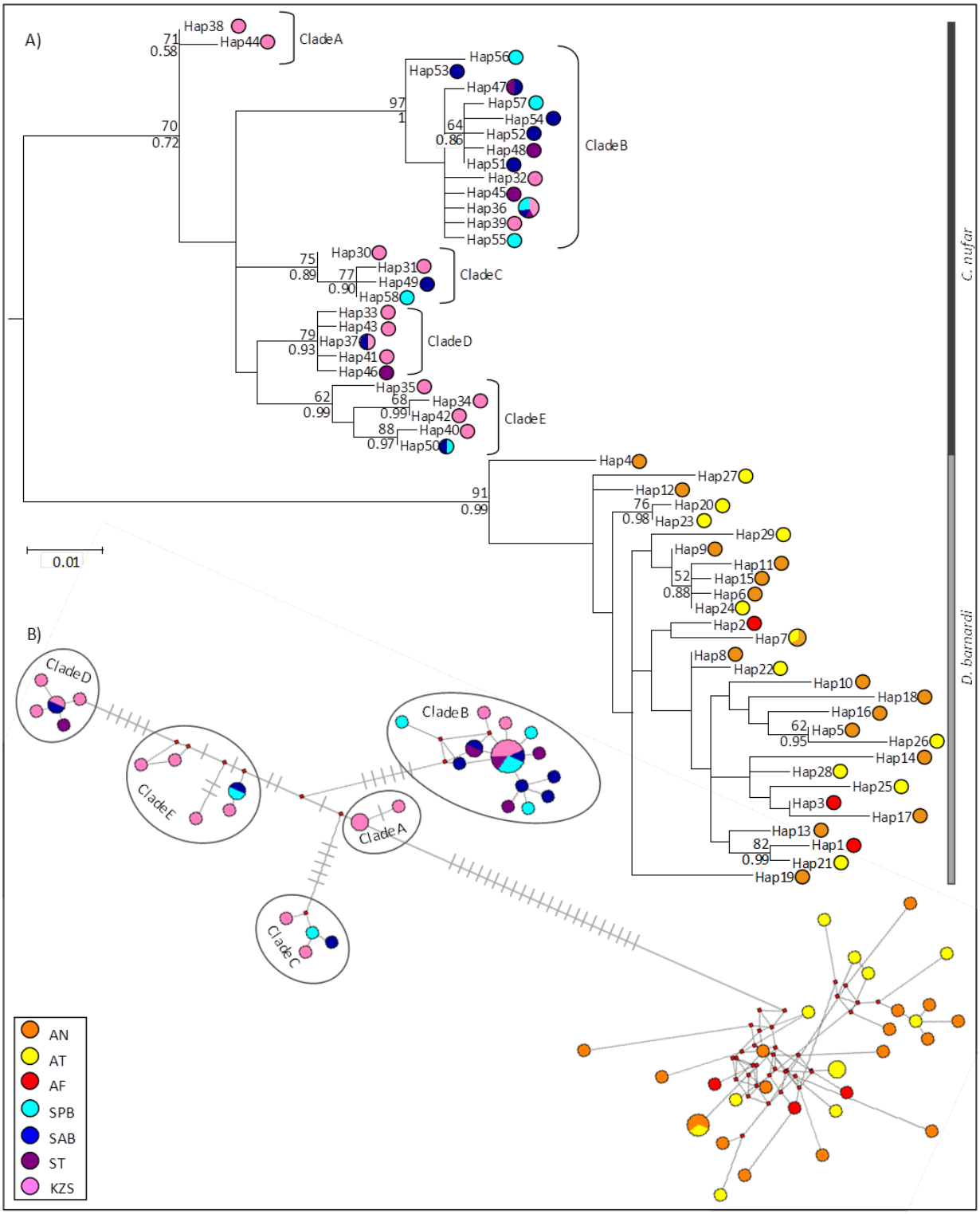


Figure 6.4: Phylogenetic relationships between and within *C. nufar* and *D. barnardi* based on 941bp of concatenated COI & CR: A) ML phylogeny (statistical support for nodes is based on Bayesian analysis (BI - below branches) and maximum likelihood bootstrap support (BS) - above branches; B) Median joining haplotype network, where branch lengths are proportional to the number of differences (each dash represents a single mutation) and node size is proportional to the haplotype frequency. Colours correspond to sampling location.

Table 6.4: Pairwise mtDNA Φ_{ST} values between samples based on 543bp of COI (below the diagonal) and 398bp CR (above the diagonal). Significant values denoted in bold.

	TFE	AN	AT	AF	SPB	SAB	ST	KZS
TFE		**0.857**	**0.876**	**0.909**	**0.892**	**0.886**	**0.907**	**0.855**
AN	**0.988**		-0.022	-0.048	**0.661**	**0.659**	**0.678**	**0.629**
AT	**0.989**	-0.003		-0.083	**0.684**	**0.681**	**0.705**	**0.645**
AF	**0.990**	0.056	-0.017		**0.709**	**0.701**	**0.757**	**0.646**
SPB	**0.986**	**0.864**	**0.860**	**0.806**		-0.098	-0.078	0.089
SAB	**0.986**	**0.865**	**0.862**	**0.820**	-0.099		-0.095	0.060
ST	**0.990**	**0.885**	**0.887**	**0.859**	-0.081	-0.075		0.127
KZS	**0.986**	**0.860**	**0.858**	**0.828**	0.107	0.090	0.188	

6.3.3 Timing of divergence (speciation) of Dentex species

To improve sequence diversity COI sequences generated in the present study were combined with those of Viret *et al.* (2018), who included a more comprehensive geographic sampling of *D. dentex* and other Sparidae species, including an unidentified cryptic lineage previously identified as *D. dentex* by Bargelloni *et al.* (2003). This unknown Sparidae lineage (labelled "Sparidae spp." in Fig. 6.5) was sampled from the north east Atlantic and clustered more closely with *C. nufar* and *D. barnardi* than with *D. dentex*. Timetree (Fig. 6.5) results placed the divergence of *D. dentex* from the Sparidae spp. / *D. barnadi* / *C. nufar* group at 14.62 MYA (95%CI = 1.457-33.767), whereas the north east Atlantic Sparidae spp. diverged from *C. nufar* and *D. barnardi* at 5.09MYA (95%CI = 0.00-13.405). Divergence of South African *C. nufar* and Angolan *D. barnardi* was placed much more recently at 1.26MYA (95%CI = 0.000-3.778).

6.4 Discussion

In line with the findings of previous phylogenetic investigations (Chiba *et al.* 2009; Sanitini *et al.* 2014; Carpenter & Johnson 2001) the mtDNA phylogenies reported here support the distinct species status of *D. dentex*, *D. barnardi* and *C. nufar*. However, contrary to expectations based on these previous Sparid phylogenies, and in support of the results of Viret *et al.* (2018), mtDNA tree topologies do not support the status of *D. dentex* as the north east Atlantic/Mediterranean sister species of *C. nufar* and *D. barnardi*. Instead, it would appear that the north east Atlantic/Mediterranean sister species of *C. nufar* and *D. barnardi* is represented by the unknown Sparid lineage resolved in the Portuguese samples of Viret *et al.* (2018), which had been previously misidentified as a divergent Atlantic *D. dentex* clade by Bargelloni *et al.* (2003). Sequence divergence support the species status of this linage, distinct from its nearest relatives (*D. barnardi* and *C. nufar*); and as such indicate a case of sympatric cryptic species

Figure 6.5: Chronogram showing estimated divergence times within the Sparidae, calculated using the RelTime method applied to the ML topology generated using 533bp of COI sequences generated in this study combined with those used in Viret *et al.* (2018). Estimated node ages (MYA) are given above the node, bootstrap support given below the node, with bars representing the 95% confidence interval of age estimates. Timescale, in million years (MY) is provided on the x axis.

(similar to that of chapters 5.1 & 5.2) in the north east Atlantic, which calls for more rigorous scientific investigation as to the distributional limits of *D. dentex* and this unknown lineage, in order to define regions of potential indiscriminate harvesting which could compromise sustainability of north east Atlantic Dentex fisheries (Griffiths & Heemstra 1995; Garcia-Vasquez *et al.* 2012). In light of this finding we discuss the eco-evolutionary modality of this species group (e.g. the unknown Sparid lineage, *D. barnardi* and *D. dentex*) across the eastern Atlantic, in relation to major geological and climatic events, and in doing so hope to build upon the work of Viret *et al.* (2018) to get a broader insight into the evolutionary history of eastern Atlantic Sparids.

Based on the phylogeny reported here diversification of the Dentex species complex roughly aligns with an evolutionary scenario that appears to be common across the eastern Atlantic (e.g. Chapters 2.1 & 3.2; Gwilliam 2017; Henriques 2012; Grant & Leslie 2001). With genetic diversity partitioned into distinct and highly divergent north east Atlantic/Mediterranean (*D. dentex* & unknown Sparid species), tropical west African (*D. barnardi*) and southern African (*C. nufar*) genetic groups, and similar to many previous investigations (e.g. Grant & Leslie 2001; Henriques *et al.* 2014a;b; Reid *et al.* 2016) the Mediterranean/north east Atlantic genetic group represents the basal position in the phylogeny, indicating either vicariance driven isolation of a once broadly distributed eastern Atlantic clade or historical dispersal and founder events southward across the eastern Atlantic as being the predominant driver of diversification in this region.

The estimated timing of divergence (~5.09 MYA) of the north-east Atlantic unknow Sparid species from its southern Atlantic sibling species (*C. nufar* and *D. barnardi*) aligns with a scenario of vicariance driven speciation in response upwelling intensification within the Canary Current Upwelling System (CCUS), as opposed to far earlier vicariant divergence in response to the formation of steep latitudinal thermal gradients in the tropics(~14.8MYA) or more recent (~2.7MYA-400KYA) founder event driven speciation, driven by trans-equatorial dispersal events during periods of global cooling (e.g. Palero *et al.* 2009; Grant & Leslie 2001). Numerous intermittent periods of upwelling intensification have been observed within the CCUS, with the most pronounced being during the Mid-Miocene Climatic Transition (MMCT- 14.2-12.8 MYA- Emery & Uchupi 1984) and the onset of global glaciation in the late Pliocene (~3MYA- Marlow *et al.* 2000). However, for all time estimates presented here 95% confidence intervals were broad, as such and in the absence of more accurate divergence date estimates, we can only make a broad inference regarding the role of the CCUS in driving diversification of these taxa which cannot be confidently associated with specific events in the history of this

upwelling system.

A comparatively more recent divergence was resolved between the tropical/subtropical Atlantic *D. barnardi* and the tropical Indian Ocean *C. nufar* (~1.26MYA). Its is likely, as has been observed for numerous species groups (Turpie *et al.* 2000; Lessios *et al.* 2003; Floeter *et al.* 2008) that the cold upwelled water of the BUS has driven speciation between the Atlantic and Indian Oceans. However whether this is due to vicariance (i.e. associated with the initial formation of the Lüderitz upwelling cell ~2.9-1.2MYA-Marlow 2000; Robinson *et al.* 2002; Krammer *et al.* 2006) or due to founder events (from South Africa to Angola or vice versa) in association with fluctuations in upwelling intensity during Pleistocene glacial/ interglacial cycles cannot be determined confidently from the dates resolved here. Although a significant proportion of species level divergences across the BUS have been attributed to a vicariance scenario (i.e. Chapter 2.1; Chapter 3.2, Henriques *et al.* 2014; 2016; Gwilliam 2017), intermittent permeability to the BUS has been reported (Chapter 2.3; 3.2; Reid *et al.* 2016; Gwilliam 2017), which alongside the star shaped haplotype networks and neutrality tests of *D. barnardi* (which support a bottleneck and expansion), support the role of founder dispersal from South Africa to Angola as the driver of speciation between *C. nufar* and *D. barnardi*.

The results reported here highlight the dominant roles of the BUS and the CCUS in limiting dispersal and ultimately driving speciation amongst these commercial important Sparid species. However, to appropriately disentangle the relative roles of vicariance and founder events in driving divergence across the eastern Atlantic a more comprehensive genetic/ genomic analysis of these species would be required, and would benefit from the inclusion of a greater set of genomic/genetic markers across a broader sampling range which could facilitate more rigorous tests of the interaction between local demographics and the timings of divergence events with the oceanographic and climatic history of the BUS and CCUS.

CHAPTER 7

Discussion

CHAPTER 7

Discussion

Through investigating the spatial distribution of genetic diversity this thesis aimed to assess the roles of ecological, environmental and climatic factors in shaping microevolution, adaptation and speciation in a number of marine taxa. From a geographical perspective the thesis had a clear focus on two regions, the east Atlantic and the South West Indian Ocean (SWIO). From the outset these regions represented very different model systems to understand the interplay between ecological and evolutionary dynamics. Specifically, the east Atlantic is punctuated by a number of prominent biogeographic barriers suggesting that genetic structuring would be largely driven by factors extrinsic to species (e.g. environmental). In contrast, the SWIO has been shown to be a high gene flow region suggesting that structuring among species with large dispersal potentials would generally be shallower and driven by intrinsic factors such as behaviours. For the relevant studies in both regions there was a cross species element providing further resolution of the roles of ecological determinism and physical structuring. In addition to understanding eco-evolutionary diversification processes an additional aim of this research was to contribute to management and conservation through: (i) delineation of management units, (ii) provision of genetic resources, and (iii) directing further studies. In this general discussion we place our findings within the broader context of ongoing and future challenges to biodiversity conservation and management in these regions.

7.1 Divergence and demographic history across the Benguela Upwelling System

The BUS is recognised as a major marine biogeographic breakpoint, halting the dispersal of taxa between the Atlantic and Indian Oceans and often initiating species level divergence events between Atlantic and Indian Ocean taxa (Turpie *et al.* 2000; Lessios *et al.* 2003; Floeter *et al.* 2008). The results reported here across multiple species and species groups support this, with most species (e.g. *Sepia officinalis* species complex- Chapter 2.3; *Dentex dentex* species complex- Chapter 6; *Chelon dumerili*- Chapter 3.2; *Chelon tricuspidens*- Chapter 3.2) reporting a breakdown in gene flow between individuals distributed in the northern and southern Benguela subsystems. Studies reporting such restricted gene flow can be grouped into three categories based on the resolved phylogeographic patterns (Table 7.1): (i) reciprocal monophyly (e.g. *C. tricuspidens*- Chapter 3.2; *Lichia amia*-Henriques *et al*. 2012; *Diplodus* capensis-Henriques 2012; Atractoscion *aequidens*- Henriques *et al*. 2014; 2016) (ii) incomplete lineage sorting (e.g. *Diplodus hottentotus*- Gwilliam *et al*. 2018), and (iii) secondary contact due to historical permeability (e.g. Sepia species-Chapter 2.3; *C. dumerili*-Chapter 3.2;

Table 7.1: Collated results of all previous population genetic and phylogeographic investigations that included representative samples from the northern (Angola) and southern (South Africa) Benguela systems. Tmrca=Time since most recent common ancestor, Texp= time since expansion. 95% confidence intervals are given in parentheses. N/A indicates where tests were not conducted.

Species	Adult habitat	Egg/larval dispersal?	Genetic marker	Structure across Benguela	Monophyly across Benguela	Tmrca (MY)	Texp northern Benguela (KY)	Texp southern Benguela (KY)	Reference
Lichia amia	pelagic	pelagic/pelagic	CR	Y	Y	0.202 (0.096-0.287)	17.76 (0-44.94)	12.69 (0-31.92)	Henriques *et al.* 2012
Pomatomos saltatrix	pelagic	pelagic/pelagic	CR & msats	Y	N	0.220 (0.077-0.395)	N/A	N/A	Reid *et al.* 2016
Octopus vulgaris	benthopelagic	fixed/ pelagic	cytb	Y	Y	0.231 (0.098-0.385)	NoExp	129.31 (18.19-150.86)	DeBeer 2014
Diplodus capensis	benthopelagic	pelagic/pelagic	COI & msats	Y	Y	0.367 (0.216-0.501)	8.14 (4.82-15.55)	40.38 (0-169.38)	Henriques 2012
Sarpa salpa	benthopelagic	Pelagic/estuarine nursery	COI, CR & msats	Y	N	0.708 (0.547-0.869)	160.00 (117.00-194.00)	74.00 (44.00-115.000	Gwilliam 2017
Lithognathus mormyrus	benthopelagic	Pelagic/estuarine nursery	COI, CR & msats	Y	N	0.786 (0.650-0.922)	259.00 (180.00-315.00)	173.00 (116.00-305.00)	Gwilliam 2017
Chelon tricuspidens	Demersal-catadromous	Pelagic/estuarine nursery	COI	Y	Y	0.940 (0.000-2.290)	N/A	N/A	Chapter 3.2
Argyrosomus coronus - A. japonicus	benthopelagic	pelagic/pelagic	CR & msats	Y	Y	1.097 (0.614-1.526)	30.97 (13.60-56.54)	24.91 (16.24-43.46)	Henriques 2012; Potts *et al*. 2014a
Dentex barnardi - Cheimerius nufar	demersal	pelagic/pelagic	CR, COI & α-amy	Y	Y	1.260 (0.000-3.780)	N/A	N/A	Supplementary material
Spondyliosoma cantharus - S. ermaginatum	demersal	demersal/ demersal	COI, CR & msats	Y	Y	1.920 (1.420-2.420)	26.00 (9.00-50.00)	42.00 (4.00-76.00)	Gwilliam 2017
Atractoscion aequidens	benthopelagic	pelagic/pelagic	CR, msats	Y	Y	2.380 (1.500-4.370)	30.30 (6.33-56.99)	24.53 (1.22-82.41)	Henriques *et al.* 2014; 2016
Sepia hierredda - S. vermiculata	demersal	fixed/ direct developing	COI, cytb, NCR, SNPs	Y (only at SNPs)	Not at mtDNA	2.710 (0.000-6.040)	89.39 (59.83-279.59)	68.68 (24.61-140.32)	Chapter 2.2 & 2.3
Chelon dumerili	Demersal-catadromous	Pelagic/estuarine nursery	COI, α-amy	Y	N	7.310 (2.590-12.02)	N/A	N/A	Chapter 3.2
Triakis magalopterus	pelagic	ovoviviparous	CR & msats	Y	N	N/A	N/A	N/A	Soekoe 2016
Thunnus albacares	pelagic	pelagic/pelagic	SNPs	Y (only at SNPs)	Not at mtDNA	N/A	N/A	N/A	Henriques 2012; Mullins *et al.* 2018
Loligo reynaudii	pelagic	fixed/ pelagic	msat	N	N/A	N/A	N/A	N/A	Van der Vyver *et al.* 2016
Trachurus capensis	pelagic	pelagic/pelagic	COI, CR, α-amy, msats	Y (only in msats)	Not at mtDNA	N/A	N/A	N/A	Chapter 4

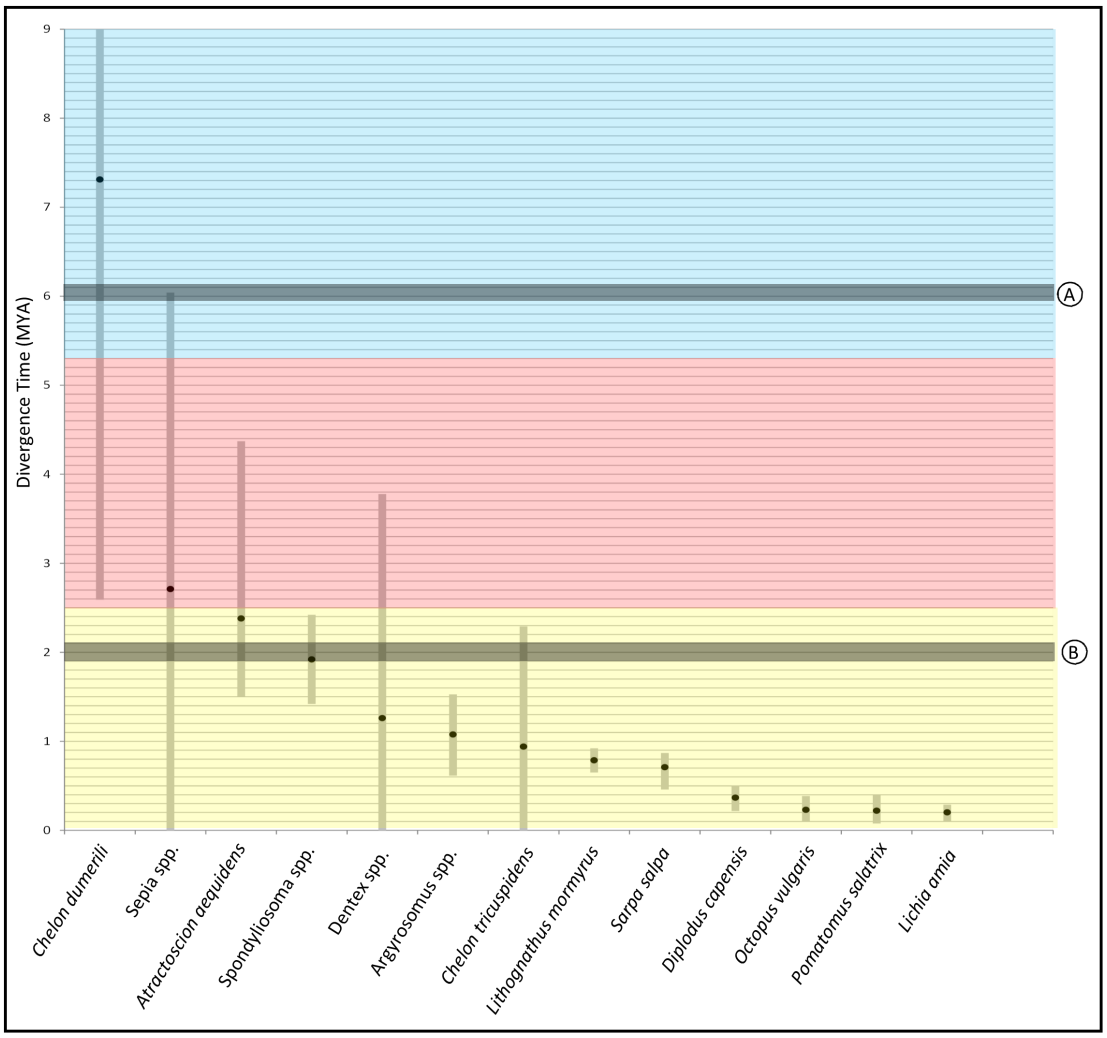


Figure 7.1: Divergence time estimates (million years ago) for various species (corresponding references can be found in Table 1) across Benguela Upwelling System. Grey bars represent 95% confidence intervals. Yellow shading corresponds to the Pleistocene epoch, red the Pliocene and blue the Miocene. Grey shading refers to major geological or oceanographic events across the eastern Atlantic: A= the establishment of south east Atlantic upwelling, B= establishment of the Lüderitz upwelling

Sarpa salpa-Gwilliam 2017; *Lithognathus mormyrus*- Gwilliam 2017 and *Pomatomus saltatrix*-Reid *et al.* 2016). Similarly dates of divergence events (Table 7.1; Fig. 7.1) associated with the BUS have generally been estimated to coincide with one of three historical oceanographic/ climatic events: (i) Pleistocene glacial-interglacial cycles (e.g. Argyrosomus species - Henriques 2012; *C. tricuspidens*- chapter 3.2; *L. mormyrus*- Gwilliam 2017; *S. salpa*- Gwilliam 2017; *D. capensis*- Henriques 2012; *P. saltatrix*- Reid *et al*. 2016; *L. amia*-Henriques *et al*. 2012 and Dentex species-Chapter 6); (ii) The formation of the Lüderitz upwelling cell (~2.9-1.2MYA-Marlow 2000; Robinson *et al*. 2002; Krammer *et al*. 2006) (e.g. Sepia species –Chapter 2.3; *A. aequidens*- Henriques *et al*. 2014; 2016; Spondyliosoma species- Gwilliam 2017); and (iii)The initial establishment of upwelling cells in the Benguela region (~7-6MYA: Diester-Haas *et al*. 2002; 1990; Krammer *et al*. 2006)(e.g. *C. dumerili*- Chapter 3.2).

The BUS has been reported as a major biogeographic breakpoint across a diverse range of

marine taxa with different life histories. This near universality and the evolutionary time frames involved obfuscates the potential roles of ecological differences among species in shaping genetic structure across the BUS. This is demonstrated when comparing the phylogenetic/population genetic partitioning of *Diplodus hottentotus* (Gwilliam *et al*. 2018) and *D. capensis* (Henriques 2012), two ecologically similar con-geners distributed across the BUS. For *D. capensis* reciprocal monophyly was reported between populations in the northern and southern Benguela system, whereas for *D. hottentotu*s genetic partitioning across the BUS were consistent with a scenario of incomplete lineage sorting. Similarly, for all species in which signals of divergence were reported across the BUS there appeared to be no obvious association with life history traits and the estimated timings of divergence. For example, within the grouping of species that were estimated to have diverged across the BUS during Pleistocene glacial-interglacial cycles, species life histories and ecologies ranged from the pelagic and highly dispersive *L. amia* and *P. saltatrix* to demersal species such as *L. mormyrus* and species with more complex catadromous life histories (*C. tricuspidens*). Likewise, the group of species in which divergence was estimated as occurring earlier in association with the formation of the Lüderitz upwelling cell consisted of the resident, demersal Sepia species (for which the direct development of eggs and larvae would limit the possibility of passive transport of larvae), and also the benthopelagic and migratory *A. aequidens* whose pelagic eggs and larval life history stages could facilitate substantial dispersal. It is therefore more likely that the different timescales of divergence reported here are an artefact of species-specific demographics, such as regional effective population sizes (which do not easily lend themselves to straightforward cross species comparisons; Lynch 2010), rather than any disparity in dispersal ability across the BUS. Furthermore for those species in which more recent divergence events were reported (e.g. Argyrosomus species - Henriques 2012; *Chelon tricuspidens*- chapter 3.2; *Lithognathus mormyrus*-Gwilliam 2017; *Sarpa salpa*- Gwilliam 2017; *Diplodus capensis*- Henriques 2012; *Pomatomus saltatrix*- Reid *et al.* 2016; *Lichia amia*- Henriques *et al.* 2012 and Dentex species-Chapter 6) a scenario in which a signal of a more ancient divergence has been lost due to the extinction of an ancestral clade cannot be ruled out.

Although the BUS is clearly a substantial biogeographic boundary region, with persistent and prolonged isolation reported between most species/species groups distributed across it, there is also evidence of historical permeability of the BUS, specifically from South Africa into Angola as observed for Sepia species (Chapter 2.3), *C. dumerili* (Chapterx) *S. salpa* (Gwilliam 2017), *L. mormyrus* (Gwilliam 2017) and *P. saltatrix* (Reid *et al.* 2016). Such asymmetric gene flow events have been variously linked to fluctuations in upwelling intensity during Pleisto-

cene glacials (Diester-Haas *et al.* 1992; Jahn *et al.* 2003) or interglacials (Mix 1986; Dean *et al.* 1984; Oberhänsli 1991).

Despite extensive phylogeographic study a fundamental consideration in comparing the estimated times of demographic events associated with the BUS is that many studies have employed different analytical approaches such as strict (e.g. Grant & Leslie 2001; Gwilliam 2017; Henriques 2012; Henriques *et al.* 2012; 2014b; DeBeer 2014; Reid *et al.* 2016) to relaxed (e.g. D'Amato *et al.* 2008; Palero *et al*. 2009) molecular clocks, and incorporated different methods of fossil, geological event and mutation rate-based calibrations. These different approaches

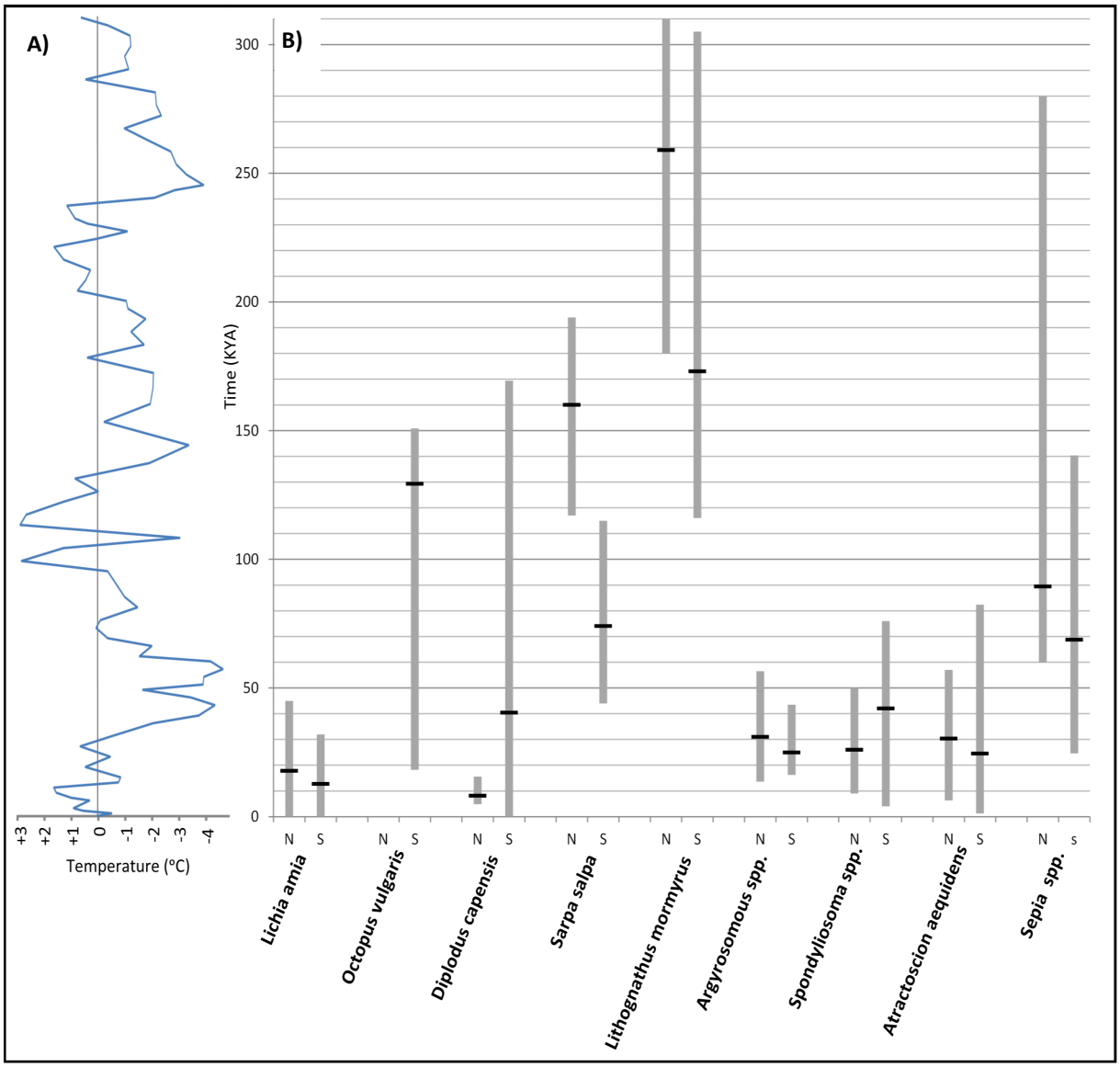


Figure 7.3: Expansion time estimated from mismatch analysis conducted for populations (references can be found in Table 1) in the northern (N) and southern (S) Benguela system. Grey bars indicate 95% confidence intervals. A) Indicates oceanic temperature fluctuations over the last 300 thousand years, relative to present day derived from cores taken from the ODP982 site. Timescales (thousand years ago) of both A and B correspond to the y axis.

each come with different caveats and potential causes for inaccuracies (reviewed in: Rutschmann 2006; Mayr 2013; Hipsley & Müller 2014;), and as such reduce comparability across studies and species. In particular, the emphasis of many studies on mutation rate-based estimates of divergence (and demographic) events across the BUS fail to incorporate the complexities of the evolutionary process, and the variability of substitution rates across even closely related and ecologically similar species (e.g. Chapter 4; Ellegren *et al.* 2003; Gillooly *et al.* 2005; Nabholz *et al.* 2007; 2008; 2009).

Demographic tests for Angolan and South African populations (Table 7.1; Fig. 7.2) indicate interactions between life history and within region demographic changes. In several benthopelagic and demersal fish (*D. capensis*- Henriques 2012; *S. salpa*- Gwilliam 2017; *L. mormyrus*- Gwilliam 2017) and cephalopod (Sepia species- see: supplementary material; *Octopus vulgaris* -DeBeer 2014) species, discordant demographic patterns for the northern and southern Benguela subsystems were resolved. Patterns ranged from signals of a bottleneck and expansion in South Africa with no such signal in Angola (*O. vulgaris*-De Beer 2014), earlier population expansions in Angola than South Africa (e.g. *S. salpa*- Gwilliam 2017; *L. mormyrus*- Gwilliam 2017, Sepia species- see: supplementary material) and vice versa (*D. capensis*- Henriques 2012). This likely reflects the bathymetric differences between Angola and South Africa and their interaction with demographic, life history and ecological traits specific to each species. Due to the presence of the Agulhas bank the continental shelf is far wider in South Africa (Siesser & Dingle 1981) than Angola and as such the rise and fall (up to 100m) of sea levels during glacial/interglacial cycles likely had profoundly different impacts on the habitats of Angola and South Africa, particularly for coastal species.

Where high gene flow/dispersal was reported across the BUS (e.g. *Trachurus capensis*-Chapter 4; *Thunnus albacares*-Mullins *et al*. 2018), these patterns can be readily associated with species life history/ecology. Although both *T. capensis* and *T. albacares* are pelagic and highly migratory (Hazin 1993; Murta *et al*. 2008), *T. capensis* is a cold adapted species, that is able to complete its life cycle in the cold waters of the BUS (Barange *et al.* 1998), whereas early life stages of *T. albacares* are less tolerant of cold waters (Brill & Lutcavage 2001; Pecoraro *et al*. 2016; 2017). Therefore, whilst the high gene flow reported amongst *T. capensis* can be aligned with the high dispersal of this species throughout the BUS, for *T. albacares* reliance on spawning sites outside of the BUS result in an absence of substantial gene flow despite high feeding dispersal across the Benguela region.

Overall the results reported in this thesis, combined with previous investigations of taxa distributed across the BUS highlight the persistence and prominence of this upwelling system as a barrier to gene flow, though there has been some historical permeability. For the vast major-

ity of species/ species groups in which the BUS was reported as a barrier to gene flow, differing patterns of genetic structuring likely reflect local demographics rather than specific life history or ecological features influencing dispersal across the BUS *per se*.

7.2 Divergence and demographic history across the south-west Indian Ocean

Similar to the eastern Atlantic, demographic analyses of SWIO taxa have reported signals of population fluctuations that can be linked to glacial/interglacial habitat shifts (Gopal *et al*. 2006; Ragionieri *et al*. 2010). However, in contrast to the west African coast, the lack of within species phylogenetic breaks across the SWIO, reported here in the mtDNA phylogenies of *Lethrinus mahsena*, *Lethrinus harak* and *Lethrinus nebulosus* (Chapters 5.1 & 5.2; Healey *et al.* 2018a; b), as well as in previous mtDNA based investigations of SWIO taxa (Ragionieri *et al.* 2009; 2010; Silva *et al.* 2010a; Hoareau *et al.* 2012; Muths *et al.* 2015), suggest a much higher gene flow throughout the SWIO, with an absence of prolonged population isolation. Across all three emperor fish species (Chapter 5.1 & 5.2; Healey *et al.* 2018a;b) genetic patterns conformed to a scenario of high gene flow with chaotic genetic patchiness. This genetic patchiness can be attributed to stochastic recruitment and/or larval cohesion (Selkoe *et al.* 2006), facilitated by the life history characteristics of these species (e.g. high fecundity and batch spawning) but also the complex oceanography of this region, dominated particularly within the Mozambique channel by a series of eddies (Swart *et al.* 2010), which are recognised as contributing to recruitment variability (Bourjea *et al.* 2007). Furthermore, unlike the BUS where clear physical structuring of species was observed, the maintenance of genetic divergence between these closely related SWIO emperor fish species (including cryptic species identified here) in sympatry suggests behavioral factors (e.g. spatial/temporal spawning migrations) and/or selection againts hybrids as isolating factors.

Paradoxical to the lack of vicariance driven diversification reported within SWIO species, a salient feature of the emperor fish datasets was the identification of cryptic diversity within what was being harvested and classified as *L. mahsena* (Chapter 5.1; Healey *et al.* 2018a) and *L. nebulosus* (Chapter 5.2; Healey *et al.* 2018b), with congruent cyto-nuclear partitioning of samples into two highly divergent and reciprocally monophyletic groups that exceeded conventional thresholds for phylogenetic species designation (De Queiroz 1998; 2007).The discovery of these apparent cryptic species, restricted to the peripheries of *L. mahsena's* and *L. nebulousus'* distribution in the Seychelles and South Africa respectively, indicates that despite a lack of vicariance in the SWIO, the large geographic distances amongst locations and/or the complex

oceanography of the region has driven speciation at the peripheries of these species' optimal habitat (Lessios *et al.* 2001; Petren *et al.* 2005).

7.3 Genetics and fisheries management of African coastal species

The data reported here adds to our understanding of historical and recurrent dispersal processes and diversification, as well as the spatial distribution of genetic diversity across the African coastline. This information pertaining to the biology, ecology, evolution and demographic history of a species is fundamental to the successful and appropriate management and conservation of marine resources (Carvalho & Hauser 1995; Reiss *et al.* 2009), as it aids in the successful alignment of management and biological units. This is of particular pertinence considering the uncertain future of growing anthropogenically driven environmental changes and exploitation that marine populations are predicted to face (Moritz 2002).

Broadly, the thesis shines a light on the limited biological investigation of African fauna when compared with other geographical regions. This is exemplified in Chapters 5.1 & 5.2, where cryptic species were resolved within SWIO populations of two commercially important and hitherto considered highly distinct fish species (*L. mahsena* and *L. nebulosus*). Such species misidentification may result in inappropriate management measures due to inaccurate estimations of biological parameters and stock abundance (Griffiths & Heemstra 1995; Garcia-Vazquez *et al.* 2012), and as a result severely compromise the sustainability of stocks. Furthermore, the identification of stochastic recruitment in the emperor fish species investigated here has the potential to not only decrease the resilience of stocks to fishing pressures, but also adds an element of unpredictability in the recovery of overexploited stocks (Kuparinen *et al.* 2014).The sustainability and persistence of SWIO fisheries is imperative given the importance of marine fisheries to food security and economy of eastern African countries, however it is estimated that 25% of western Indian Ocean fisheries are overexploited, with the remaining 75% currently exploited at the maximum limits of biological productivity (Van der Elst *et al.* 2005; Kimani *et al.* 2009).

Similar to the SWIO (Ridgway & Sampayo 2007) geographical blind spots have been highlighted by other researchers for South Africa and Angola (Von der Heyden 2009; Teske *et al.* 2011), and the Gulf of Guinea Large Marine Ecosystem (GGLME) (Mensah & Quaatey 2002). The coastlines of these regions are bound by developing countries, heavily reliant on fishery resources for food security and economic sustainability of their growing populations (van der Elst *et al.* 2005; Gilpin 2007). However, the collapse of African fisheries is a growing threat due to overexploitation (Mensah & Quaatey 2002; Kimani *et al.* 2009). In addition illegal, unregulated, and unreported harvesting represents a major threat to fishery sustainability, particular-

ly for African nations. Substantial poaching has been reported in the GGLME with illegal vessels estimated as exploiting around 370 million dollars' worth of unreported and unregulated marine resources from this region annually (Agnew *et al.* 2009). The species barcodes and regionally distinct lineages reported here thus offer potentially useful tools for fish traceability at both the species and regional level (Helyar *et al*. 2014).

Failure to identify local stocks may lead to local overfishing and severe decline. Due to the broad geographic scope many of the chapters in this thesis were macrogeographical in nature, with large geographic distances between sampling locations (>1000 nautical miles). However, it is recognised that in order to encapsulate finer scale structuring and identify seascape drivers of such structure sampling over a finer geographic scale is needed (Chakraborty 1992; Balloux & Lugon-Moulin 2002). Furthermore, stock assessment models used in fisheries management require precise and accurate data on the population identity of harvested fish to maximize long-term fisheries yield at minimal risk to population viability. There is increasing concern that many fish stocks are managed with insufficient knowledge on stock structure and exploitation rates (e.g. Reiss *et al*., 2009). Due to the large effective population sizes (Laconcha *et al.* 2015; Waples *et al.* 2018) and great propensity for dispersal of many marine species, signals of fishery relevant structuring may be beyond the resolution of traditional population genetic methods that employ a small number of putatively neutral loci. Instead, and where possible such studies would benefit from employing genomic assays (e.g. RADseq), to increase the resolution and power of neutral population genetic analyses (e.g. Mamoozadeh *et al.* 2019), as well as provide novel tools to assess non-neutral and adaptive structure within species. These markers under selection have proven powerful tools in the delineation of management units, particularly in pelagic systems (e.g. Mullins *et al.* 2018; Nielsen *et al.* 2012), furthermore these loci are expected to be associated with fitness-related traits (e.g. disease resistance- Savage & Zamudio 2016) and as such represent key tools in understanding the adaptive potential of a species or population under future environmental changes (Allendorf *et al.* 2010).

7.4 Genetics as a tool for disentangling species resilience under changing climates.

Widespread environmental changes including climate change and selective harvesting of wild populations via fishing are affecting both species distributions and populations dynamics (Coleman & Williams 2002; Burrows *et al.* 2011; 2014), with profound consequences for the sustainability of fisheries. Populations may respond to environmental change in 3 ways. Firstly, in the short term populations may acclimate to changing conditions via phenotypic plasticity. Plasticity permits rapid responses to environmental change, however, if phenotypic changes

are mostly due to plasticity, responses to continued environmental change will be limited (Gienapp *et al*. 2008). Second, populations may become adapted to conditions through genetic selection. Such adaptation requires longer time frames but allows for response to a continually changing environment beyond the limits of phenotypic plasticity (Davis *et al*. 2005; Chevin *et al*. 2010). However, if environmental stress reverses, such as in situations where fishery policies are amended, plastic changes could reverse rapidly while genetic adaptations would be more resistant. Plasticity and adaptation may also underpin the third response which is for populations to track their optimal environment in space (habitat tracking-e.g. Burrows *et al.* 2011; 2014).

In addition to the "move, adapt, acclimate, or die" categories of species' responses to climate change another potential outcome is that of species collapse via interspecific hybridisation (e.g. Valema *et al.* 2012; Taylor *et al.* 2006; Seehausen *et al.* 1997). Such interspecific permeability has been demonstrated in chapter 2.3 (for the *S. officinalis* species complex) and could be a factor in cases where: (i) formerly vicariant groups come into contact due to range shifts and intrinsic reproductive isolating barriers are incomplete (e.g. Bernatchez *et al*. 1995; Doiron *et al.* 2002; Lamaze *et al.* 2012) or (ii) where reproductive isolation is maintained by ecologically divergent selection and such selection regimes are disrupted by environmental change (e.g. pollution – Seehausen *et al*. 1997; Vonlanthen *et al*. 2012).

7.4.1.South Africa and Angola as key areas of vulnerability under changing climates: integrating patterns of genetic connectivity with species distribution models.

Patterns of genetic structuring resolved in this thesis add to evidence that many species/ populations in Angola and South Africa are trapped by regional biogeographic barriers, with growing evidence that these barriers are going to become even stronger. The intensity of upwelling currents are expected to increase under climate change due to changes in land –sea pressure and wind patterns (Snyder *et al.* 2003; Bakun *et al.* 2010; Narayan *et al.* 2010), leading to a strengthening of the BUS biogeographic boundary. Furthermore, the interaction of the cooling waters of the BUS, with climate change driven warming of tropical waters at its northern (Lima & Wethey 2012; Potts *et al.* 2014a) and southern (Rouault *et al.* 2010; Rhein *et al.* 2013) peripheries are predicted to steepen thermal gradients and cause a marked reduction in the extent of warm temperate habitat in Angola and South Africa.

The coast of Angola is predicted to warm dramatically in oncoming decades, with this warming predicted to be particularly pronounced in southern Angola, a region which has been identified as an ocean warming hotspot with SST between 1982 to 2009 increasing by around 0.8°C per decade (Lima & Wethey 2012; Potts *et al.* 2014a). Under these warming conditions tem-

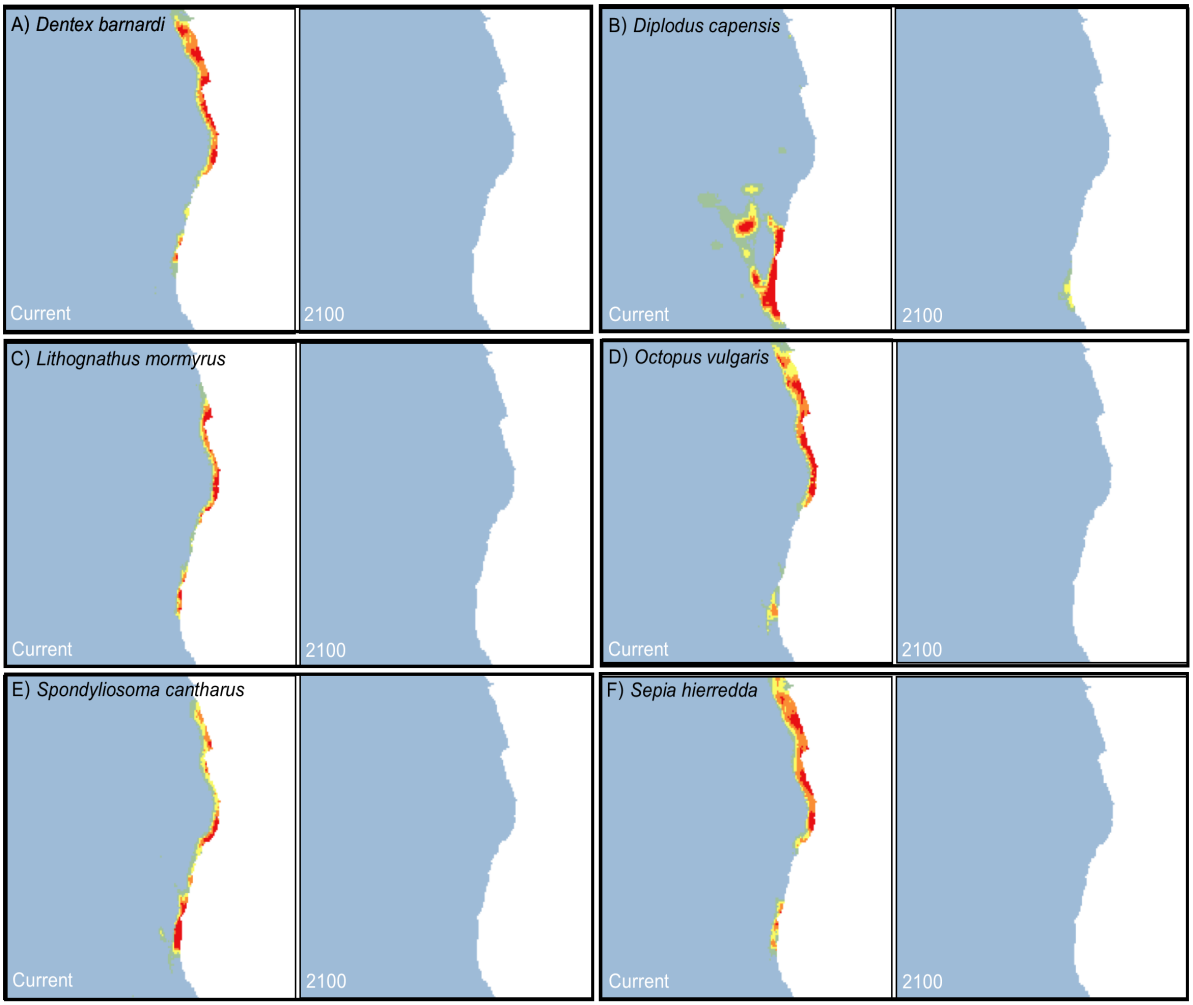


Figure 7.3: Results of Ecological niche models representing areas of predicted suitable habitat for: a) *Dentex barnardi*, b) Diplodus capensis, c) *Lithognathus mormyrus*, d) *Octopus vulgaris* e) *Spondyliosoma cantharus*, f) *Sepia hierredda* across the Angolan coast. Maps represent optimum environmental niches modelled using either environmental variables representative of current environmental conditions and also forecasted using environmental variables predicted under the RCP85 climate change scenario for the year 2100.

perature sensitive species residing across the Angolan coast can either adapt to or shift their distribution in response to SST rises. However, the poleward range shift that would theoretically be required to track optimal thermal niches is restricted by the cold BUS, a recognised boundary to dispersal for temperate species. Seven species (*Sepia hierredda*-Chapter 2.3; *Dentex barnardi*- Chapter 6 *Diplodus capensis*- Henriques 2012; *Lithognathus mormyrus*, *Sarpa salpa*, *Spondyliosoma cantharus*- Gwilliam 2017; *Octopus vulgaris*- DeBeer 2014) were found to harbour genetic diversity in Angolan waters that was restricted from and unique to genetic diversity reported in regions outside of Angola, indicating a distinct lack of recurrent dispersal outside of Angola (to both the north across the tropics of west Africa and to the south across the BUS). For these species, SDMs (implemented using methodology explained in Supplementary data: 7.2) reported a total loss of suitable habitat along the Angolan coastline, which was predicted for all species investigated (Fig. 7.3), which could have severe biodiversity implica-

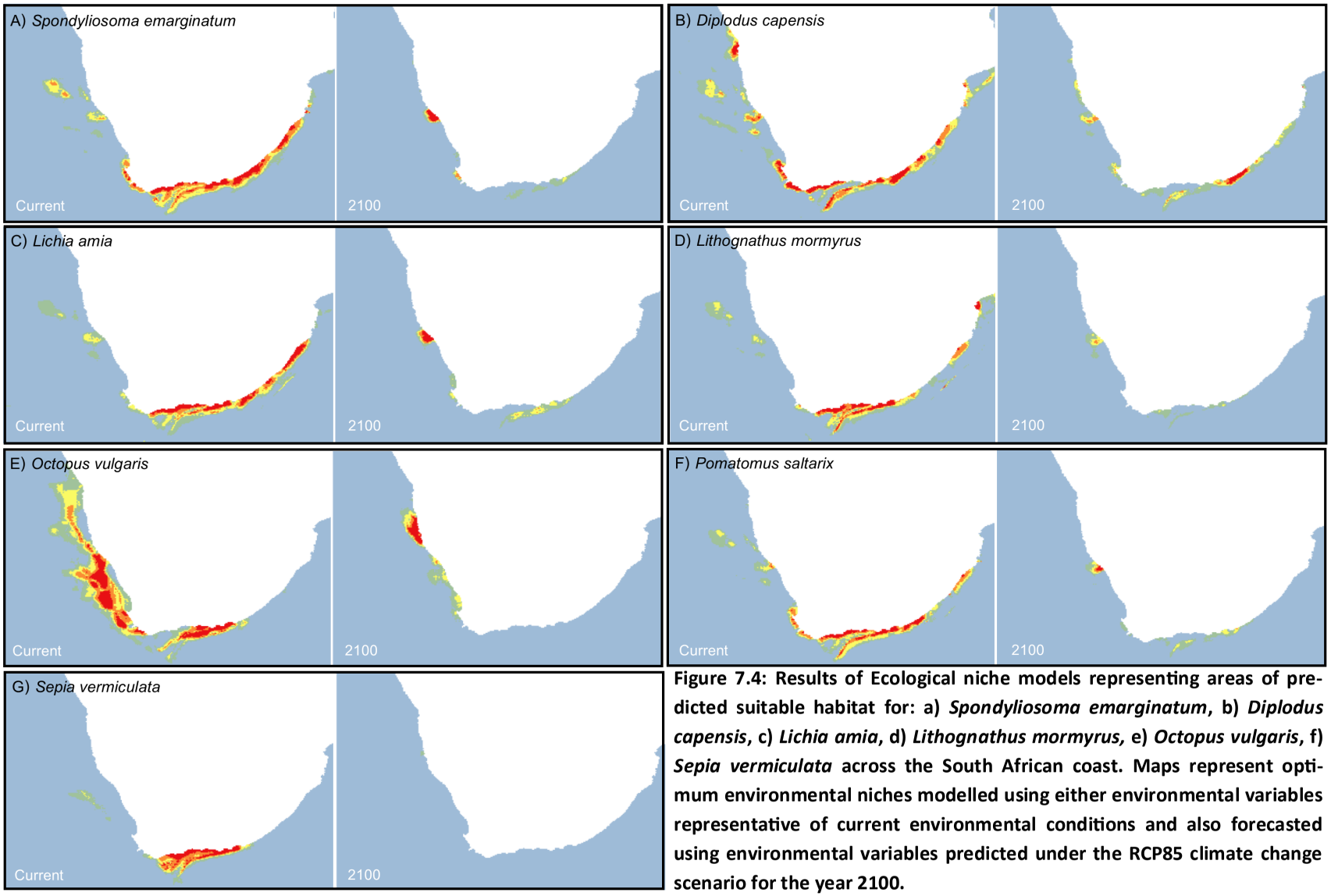


Figure 7.4: Results of Ecological niche models representing areas of predicted suitable habitat for: a) *Spondyliosoma emarginatum*, b) *Diplodus capensis*, c) *Lichia amia*, d) *Lithognathus mormyrus,* e) *Octopus vulgaris*, f) *Sepia vermiculata* across the South African coast. Maps represent optimum environmental niches modelled using either environmental variables representative of current environmental conditions and also forecasted using environmental variables predicted under the RCP85 climate change scenario for the year 2100.

tions for Angolan taxa, and as such necessitates further investigation and monitoring regarding the vulnerability of this region under climate change.

Since 1981 the coastline of South Africa has been warming at a rate comparable to global averages (Rouault *et al.* 2010), however climate change driven strengthening of oceanic wind has been predicted to increase the intensity of the cold BUS (Rouault *et al.* 2010; Lima & Wethey 2012; Lough 2012) to the west and the tropical Agulhas current (Rhein *et al.* 2013) to the east, causing an influx and convergence of cold upwelled and tropical waters along the South African coast and promoting anomalous SST cooling on the western South African coast, whilst the eastern coast is predicted to warm dramatically (Lough 2012), potentially steepening thermal gradients and contracting the extent of temperate waters along the South African coast. For temperate taxa that currently reside on the western and southern coasts of South Africa, the poleward range shifts required to theoretically track optimum thermal niches in a changing climate are obstructed due to the large expanse of open ocean separating South Africa and the nearest body of land to the South, making the endemic diversity of South Africa particularly vulnerable to changing climates. For those species where genetic/distribution data reported isolation of temperate South African populations/species from those in the eastern Atlantic and Indian Ocean (*Spondyliosoma emarginatum*-Gwilliam 2017; *Diplodus capensis*- Henriques 2012; *Lichia amia*-Henriques *et al.* 2012, *Lithognathus mormyrus*- Gwilliam 2017; *Octopus vulgaris*-DeBeer 2014; *Pomatomus saltatrix*- Reid *et al.* 2016; *Sepia vermiculata*-Chapters 2.2 & 2.3; Healey *et al.* 2017), SDMs (Fig. 7.4) identified a dramatic reduction in areas of suitable habitat for all species under various climate change scenarios, alongside substantial habitat fragmentation (with several notable refuge areas identified along the western South African coat) and in the case of *S. vermiculata* complete loss of suitable habitat across the South African coast.

Although the predictions of these SDMs highlight key areas of vulnerability under climate change, without a more thorough appreciation of the species' physical limitations, plasticity and adaptability (Pörtner & Knust 2007; King *et al.* 2017; 2019) an accurate reflection of their distributional response to environmental changes and the precise mechanisms (e.g. oxygen limitation-Pörtner & Knust 2007; spawning cues - Pankhurst & Munday 2011; Potts *et al.* 2014b; early life history stage survival- Hoegh-Guldberg & Bruno 2010 and fecundity-Pörtner *et al.* 2001) behind these shifts cannot be achieved. Nonetheless, in the absence of dispersal pathways that would facilitate effective habitat tracking, for a number of warm temperate species in Angolan and South African waters, persistence under environmental changes will be determined by *in situ* genetic and plastic machinery (Wernberg *et al*. 2016).

7.5 Methodological limitations

Although the data reported here contributes to a wider understanding of the evolutionary diversification of species residing on the African coast, there are several key constraints to the experimental design that limited the scope of this investigation. One of the major limitations of the present work was the sampling strategy. Due to the broad geographic scope of this investigation, in combination with financial, logistical and time constraints only key regions of many species' distributions were targeted, with large geographic distances between some sampling locations (>1000 nautical miles). However, it is recognised that in order to address many key questions in regards to the evolutionary history of a species and not disregard any intraspecific diversity or subpopulation structure, samples are required that are spaced evenly over a finer geographic scale and encompass the species' entire distributional range (Balloux & Lugon-Moulin 2002). As such a major limitation of any interpretations drawn from this study is the lack of fine scale population sampling both across each species' distribution but more specifically across the African coastline, which could have allowed the resolved genetic breaks to be more rigorously associated with the various biogeographic boundaries proposed in this region (e.g. the CCUS, west African tropics, BUS and Atlantic-Indian ocean boundary). This is particularly emphasised across the GGLME region, where although genetic breaks were resolved in several species, the broad sampling regime did not allow for these breaks to be directly associated with historical and contemporary oceanographic features and conditions.

In this thesis various combinations of molecular genetic markers were implemented to characterise patterns of inter-and intra-specific diversification across the African coastline, including mtDNA (Chapters 2-6) and nuclear (Chapters 3.2 & 6) sequences, microsatellites (Chapters 4, 5.1 & 5.2) and SNPs (Chapter 2.3), however in Chapters 3.1 & 3.2 only a single locus (mtDNA) was used for the majority of species, with no corresponding nuclear genetic corroboration of phylogeographic patterns. An important consideration when interpreting single locus analysis is that although diversification can most commonly be the result of processes such as reduced gene flow and genetic drift that would leave synonymous signatures on organellar and cellular genomes (but may be obscured by incomplete linage sorting), when processes such as recombination and selection affect the diversification process they can leave heterogeneously distributed signatures, both across and within cellular and organellar genomes (Whitlock & McCauley 1999). This decoupling of mitochondrial and nuclear genomes is highlighted in Chapter 2.3, where asymmetric nuclear introgression has promoted complete species level discordance in the phylogenetic signals resolved from each genome. Ultimately this highlights the need for corroborating data from the nuclear genome to complement any mtDNA partitioning resolved within chapters 3.1 and 3.2.

7.6 Future work

Understanding and predicting how regionally isolated populations in Angola and South Africa will respond to climate change will require disentangling such plastic and genetic components of resilience (Pörtner & Knust 2007; Hoffmann & Sgro 2011; Alvarez *et al* 2015; King *et al.* 2017; 2019). While population genomic analysis such as RADseq can provide insights into patterns of local adaptation and genetic structure (e.g. André *et al.* 2011; Ruzzante *et al.* 2011; Milano *et al.* 2014; Bekkevold *et al.* 2015), identifying genetic and plastic components of resilience will require combining genomics with common garden approaches (e.g. King *et al* 2019). For many species this may be unfeasible for adult life stages (e.g. tuna), however, studies of larval stages may be highly informative and are currently underway. Similar studies in which the relative influence of plastic and genetic components on species resilience would be informative in the case of the SWIO, as given the seeming lack of physical barriers to habitat tracking there is a greater possibility of ecosystem disruption through invasive species. Therefore, an appreciation of the role of intraspecific heterogeneity in genetic and plastic environmental interactions (e.g. Henkel & Hoffman 2008; King *et al*. 2019) will permit more accurate forecasting of invasive species.

An emerging theme in evolutionary biology is the role that alterations in genomic architecture (e.g. inversions, deletions, insertions, duplications, translocations and fusions) play in the evolutionary processes (reviewed in: Wellenreuther & Bernatchez 2018; Wellenreuther *et al.* 2019). Genomic rearrangements have been shown to play an important role in the development of reproductive isolation amongst species (Kozak *et al*. 2009; Lee *et al*. 2017; Charelsworth & Barton 2018; Hooper 2019), and have been linked to the obstruction of genetic exchange amongst sympatric lineages (Noor *et al*. 2001; Castiglia 2014; Hooper 2016). Furthermore, genomic rearrangements, particularly in the form of chromosomal inversions are increasingly associated with ecotype (e.g. sticklebacks-Jones *et al.* 2012; cichlids- Brawand *et al.* 2014; Fan & Meyer 2014; Drosophila-Fuller *et al.* 2019; Kapun & Flatt 2019; Littorina snails-Faria *et al.* 2019; but see Barth *et al.* 2019), and as such have been putatively linked to adaptation to heterogenous environments. Analysis of such rearrangements could fill knowledge gaps pertinent to our understanding of the neutral and functional diversification of eastern Atlantic taxa and may prove particularly informative in understanding the drivers and mechanisms promoting adaptation and speciation across the BUS.

7.7 Conclusions

Through the addition of novel datasets pertaining to previously unstudied taxa this thesis builds upon previous population genetic and phylogeographic investigations of the Atlantic and Indian Ocean coasts of Africa, and contributes to a wider understanding of the patterns and processes of eco-evolutionary diversification in these relatively under investigated regions. In particular the results resolved here, when combined with previous studies highlight: (i) the strength and persistence of the BUS as a biogeographic barrier to dispersal, driver of allopatric speciation and its central role in shaping Atlantic marine biodiversity; (ii) the unique genetic composition, for many taxa, of Gulf of Guinea populations, an area that is a geographical blind spot of many eastern Atlantic phylogeographic/population genetic investigations. Furthermore, the inclusion of comparative phylogeographic data from the SWIO, in which a lack of pronounced intra-specific phylogenetic breaks were resolved, highlighted the demographic openness of the SWIO when compared with the eastern Atlantic, which appears to be more heavily influenced by historical climatic and geological events. Ultimately these results have important implications for the management of African marine resources (e.g. in determining stock structure, the identification of cryptic diversity and detection of recruitment variability), and conservation of endemic ecological and genetic diversity of the African coastline.

BIBLIOGRAPHY

Abaunza, P., Gordo, L.S., Santamaría, M.G., Iversen, S.A., Murta, A.G., and Gallo, E. (2008b). Life history parameters as basis for the initial recognition of stock management units in horse mackerel (*Trachurus trachurus*). *Fisheries Research*, 89(2), 167-180.

Abaunza, P., Murta, A., Mattiucci, S., Cimmaruta, R., Nascetti, G., Magoulas, A., ...and Santos, A.T. (2004). Stock discrimination of horse mackerel (Trachurus trachurus) in the Northeast Atlantic and Mediterranean Sea: integrating the results from different stock identification approaches. ICES.

Abaunza, P., Murta, A.G., Campbell, N., Cimmaruta, R., Comesaña, A.S., Dahle, G., Santamaría, M.G., and Magoulas, A. (2008a). Stock identity of horse mackerel (*Trachurus trachurus*) in the Northeast Atlantic and Mediterranean Sea: Integrating the results from different stock identification approaches. *Fisheries Research,* 89(2), 196-209.

Abaunza, P., Gordo, L.S., Santamaría, M.G., Iversen, S.A., Murta, A.G., & Gallo, E. (2008b). Life history parameters as basis for the initial recognition of stock management units in horse mackerel (*Trachurus trachurus*). *Fisheries Research*, 89(2), 167-180.

Aboim, M.A., Menezes, G.M., Schlitt, T., and Rogers, A.D. (2005). Genetic structure and history of populations of the deep-sea fish Helicolenus dactylopterus (Delaroche, 1809) inferred from mtDNA sequence analysis. *Molecular Ecology*, 14(5), 1343-1354.

Agenbag, J.J., and Shannon, L.V. (1988). A suggested physical explanation for the existence of a biological boundary at 24 30′ S in the Benguela system. *South African Journal of Marine Science*, 6(1), 119-132.

Agnew, D.J., Pearce, J., Pramod, G., Peatman, T., Watson, R., Beddington, J.R., and Pitcher, T.J. (2009). Estimating the worldwide extent of illegal fishing. *PloS one*, 4(2), e4570.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19 (6), 716-723.

Akasaki, T., Nikaido, M., Tsuchiya, K., Segawa, S., Hasegawa, M., and Okada, N. (2006). Extensive mitochondrial gene arrangements in coleoid Cephalopoda and their phylogenetic implications. *Molecular phylogenetics and evolution*, *38(*3), 648-658.

Albaret, J.J. (1984). Premières observations sur la faune ichtyologique de la Casamance. Archive Scientifique du Crodt, 131. Dakar: Centre de Recherches Océanographiques de Dakar-Thiaroye.

Albaret, J.J. (1987). La faune ichtyologique de la Casamance: observations réalisées en 1984–1985. Document Scientifique 105. Dakar: Centre de Recherches Océanographiques de Dakar-Thiaroye.

Albaret, J.J. (2003). Mugilidae. In: Paugy ID, Lévêque C, Teugels GG, editors. Poissons d'eaux douces et saumâtres de l'Afrique de l'Ouest tome 2: Collection Faune et flore Tropicales. Tervuren, Belgium: Musée Royal de l'Afrique Centrale, and Paris: ORSTOM, pp. 780–786.

Albaret, J.J., and Legendre, M. (1985). Biology and ecology of Mugilidae in Ebrié lagoon (Ivory Coast): Potential interest for lagoon aquaculture. *Revue d'Hydrobiologie Tropicale*, 18(4), 281-303.

Albaret, J.J., Mugilidae, D.P., Lévêque, C., and Teugels, G.G. (2003). *The Fresh and Brackish Water Fishes of West Africa.* IRD Editions, Paris, 601-611.

Albert, A.Y.K., and Schluter, D. (2004). Reproductive character displacement of male stickleback mate preference: reinforcement or direct selection?. *Evolution*, 58(5), 1099-1107.

Allendorf, F.W., Hohenlohe, P.A., and Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature reviews genetics*, 11(10), 697.

Allendorf, F.W., Leary, R.F., Spruell, P., and Wenburg, J.K. (2001). The problems with hybrids: setting conservation guidelines. *Trends in ecology and evolution*, 16(11), 613-622.

Almada, V.C., Oliveira, R.F., Goncalves, E.J., Almeida, A.J., Santos, R.S., and Wirtz, P. (2001). Patterns of diversity of the north-eastern Atlantic blenniid fish fauna (Pisces: Blenniidae). *Global Ecology and Biogeography*, 10(4), 411-422.

Alter, S.E., and Palumbi, S.R. (2009). Comparing evolutionary patterns and variability in the mitochondrial control region and cytochrome b in three species of baleen whales. *Journal of Molecular Evolution*, *68*(1), 97-111.

Alvardo-Bremer, J.R., Viñas, J., Mejuto, J., Ely, B., and Pla, C. (2005). Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. *Molecular phylogenetics and evolution*, 36(1), 169-187.

Alvarez, M., Schrey, A.W., and Richards, C.L. (2015). Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution?. *Molecular ecology*, 24(4), 710-725.

Alves, M.L., and Colin de Verdière, A. (1999). Instability dynamics of a subtropical jet and applications to the Azores Front Current System: eddy-driven mean flow. *Journal of Physical Oceanography*, 29(5), 837-864.

Amor, M.D., Laptikhovsky, V., Norman, M.D., and Strugnell, J.M. (2015). Genetic evidence extends the known distribution of *Octopus insularis* to the mid-Atlantic islands Ascension and St Helena. *Journal of the Marine Biological Association of the United Kingdom*: 1–6.

Anderson, S.C., Flemming, J.M., Watson, R., and Lotze, H.K. (2011). Rapid global expansion of invertebrate fisheries: trends, drivers, and ecosystem effects. *PLoS ONE,* 6: e14735.

André, C., Larsson, L.C., Laikre, L., Bekkevold, D., Brigham, J., Carvalho, G.R., ... and Ruzzante, D.E. (2011). Detecting population structure in a high gene-flow species, Atlantic herring (Clupea harengus): direct, simultaneous evaluation of neutral vs putatively selected loci. *Heredity*, 106(2), 270.

Aoki, M., Imai, H., Naruse, T., and Ikeda, Y. (2008). Low genetic diversity of oval squid, Sepioteuthis cf. lessoniana (Cephalopoda: Loliginidae), in Japanese waters inferred from a mitochondrial DNA non-coding region. *Pacific Science*, 62(3), 403-411.

Arhan, M., Colin De Verdiére, A., and Mémery, L. (1994). The eastern boundary of the subtropical North Atlantic. *Journal of Physical Oceanography*, 24(6), 1295-1316.

Arkhipkin, A.I. (2005). Statoliths as 'black boxes'(life recorders) in squid. *Marine and freshwater research*, 56 (5), 573-583.

Arkhipkin, A.I., Rodhouse, P G., Pierce, G.J., Sauer, W., Sakai, M., Allcock, L., ... and Chen, C.S. (2015). World squid fisheries. *Reviews in Fisheries Science and Aquaculture*, 23(2), 92-252.

Arndt, A., and Smith, M.J. (1998). Genetic diversity and population structure in two species of sea cucumber: differing patterns according to mode of development. *Molecular Ecology*, 7(8), 1053-1064.

Atarhouch, T., Rami, M., Naciri, M., and Dakkak, A. (2007). Genetic population structure of sardine (Sardina pilchardus) off Morocco detected with intron polymorphism (EPIC-PCR). *Marine Biology*, 150(3), 521-528.

Atarhouch, T., Rüber, L., Gonzalez, E.G., Albert, E.M., Rami, M., Dakkak, A., and Zardoya, R. (2006). Signature of an early genetic bottleneck in a population of Moroccan sardines (Sardina pilchardus). *Molecular Phylogenetics and Evolution*, 39(2), 373-383.

Atkinson, L.J., Leslie, R.W., Field, J.G., and Jarre, A. (2011). Changes in demersal fish assemblages on the west coast of South Africa, 1986–2009. *African Journal of Marine Science*, 33(1), 157-170.

Autem, M., and Bonhomme, F. (1980). Eléments de systématique biochimique chez les Mugilidés de Méditerranée. *Biochemical Systematics and Ecology*, 8(3), 305-308.

Avise, J.C. (2000) Phylogeography. Harvard University Press, Cambridge, Mass.

Avise, J.C., Helfman, G.S., Saunders, N.C., & Hales, L.S. (1986). Mitochondrial DNA differentiation in North Atlantic eels: population genetic consequences of an unusual life history pattern. *Proceedings of the National Academy of Sciences*, 83(12), 4350-4354.

Avise, J.C., and Saunders, N.C. (1984). Hybridization and introgression among species of sunfish (Lepomis): analysis by mitochondrial DNA and allozyme markers. *Genetics*, 108(1), 237.

Axelsen, B.E., Krakstad, J.O., and Bauleth-D'Almeida, G. (2004). Aggregation dynamics and diel vertical migration in Cape horse mackerel (*Trachurus trachurus capensis*) in the Namibian Benguela. Implications for acoustic abundance estimation. In U. R. Sumaila, S. I. Steinshamn, M. D. Skogen and D. Boyer (Eds.), *Ecological, Economic and Social Aspects of Namibian Fisheries* (pp. 135-164). Delft, Netherlands: Eburon Publishers.

Bacheler, N.M., Wong, R.A., and Buckel, J.A. (2005). Movements and mortality rates of striped mullet in North Carolina. *North American Journal of Fisheries Management*, 25(1), 361-373.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., ... and Johnson, E.A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one*, 3(10), e3376.

Bakun, A., Field, D.B., Redondo-Rodriguez, A.N.A., and Weeks, S.J. (2010). Greenhouse gas, upwelling-favorable winds, and the future of coastal ocean upwelling ecosystems. *Global Change Biology*, 16(4), 1213-1228.

Balloux, F., and Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Molecular ecology*, 11(2), 155-165.

Bandelt, H.J., Forster, P., and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular biology and evolution*, 16(1): 37–48.

Barange, M., and Boyd, A.J. (1992). Life history, circulation and maintenance of Nyctiphanes capensis (Euphausiacea) in the northern Benguela upwelling system. *South African Journal of Marine Science*, 12(1), 95-106.

Barange, M., Pillar, S.C., and Hampton, I. (1998). Distribution patterns, stock size and life-history strategies of Cape horse mackerel *Trachurus trachurus capensis*, based on bottom trawl and acoustic surveys. *African Journal of Marine Science*, 19.

Barber, P.H., Erdmann, M.V., and Palumbi, S.R. (2006). Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the coral triangle. *Evolution*, 60 (9), 1825-1839.

Bargelloni, L., Alarcon, J.A., Alvarez, M.C., Penzo, E., Magoulas, A., Reis, C., and Patarnello, T. (2003). Discord in the family Sparidae (Teleostei): divergent phylogeographical patterns across the Atlantic–Mediterranean divide. *Journal of evolutionary biology*, 16(6), 1149-1158.

Bargelloni, L., Alarcon, J.A., Alvarez, M.C., Penzo, E., Magoulas, A., Palma, J., and Patarnello, T. (2005). The Atlantic–Mediterranean transition: discordant genetic patterns in two seabream species, Diplodus puntazzo (Cetti) and Diplodus sargus (L.). *Molecular phylogenetics and evolution*, 36(3), 523-535.

Barraclough, T.G., and Vogler, A.P. (2000). Detecting the geographical pattern of speciation from species-level phylogenies. *The American Naturalist*, 155(4), 419-434.

Barraclough, T.G., Vogler, A.P., and Harvey, P.H. (1998). Revealing the factors that promote speciation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 353(1366), 241-249.

Barth, J.M., Damerau, M., Matschiner, M., Jentoft, S., and Hanel, R. (2017). Genomic differentiation and demographic histories of Atlantic and Indo-Pacific yellowfin tuna (Thunnus albacares) populations. *Genome biology and evolution*, 9(4), 1084-1098.

Barth, J.M., Villegas-Ríos, D., Freitas, C., Moland, E., Star, B., André, C., ... and Righton, D. (2019). Disentangling structural genomic and behavioural barriers in a sea of connectivity. *Molecular ecology*. 28(6), 1394-1411.

Barton, E.D., Arıstegui, J., Tett, P., Cantón, M., Garcıa-Braun, J., Hernández-León, S., ... and Basterretxea, G. (1998). The transition zone of the Canary Current upwelling region. *Progress in Oceanography*, 41(4), 455-504.

Bat, L., Erdem, Y., Ustaoğlu, S., Yardım, Ö., & Satılmış, H.H. (2005). A study on the fishes of the Central Black Sea coast of Turkey.

Battistuzzi, F.U., Tao, Q., Jones, L., Tamura, K., and Kumar, S. (2018). RelTime relaxes the strict molecular clock throughout the phylogeny. *Genome biology and evolution*. 10(6), 1631-1636.

Bauchot, M.L., Hureau, J.C., (1986). Fishes of the North-eastern Atlantic and the Mediterranean. *Vol. II. UNESCO. Fish. Ne. Atl. and Mediterranean*, 517-1007.

Bauchot, M.L., Smith, M.M., Sparidae, W.F., & Bianchi, W. (1984). FAO species identification sheets for fishery purposes. *Western Indian Ocean (Fishing Area 51). Rome*.

Beekman, M., Dowling, D.K., and Aanen, D.K. (2014). The costs of being male: are there sex-specific effects of uniparental mitochondrial inheritance?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1646), 20130440.

Beheregaray, L.B. (2001). *Molecular evolution, biogeography and speciation of the Neotropical fish genus Odontesthes*. Macquarie University.

Beheregaray, L.B. (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, 17, 3754-3774.

Beheregaray, L.B., and Sunnucks, P. (2001). Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish Odontesthes argentinensis. *Molecular Ecology*, 10(12), 2849-2866.

Bekkevold, D., Helyar, S.J., Limborg, M.T., Nielsen, E.E., Hemmer-Hansen, J., Clausen, L.A., and Carvalho, G.R. (2015). Gene-associated markers can assign origin in a weakly structured fish, Atlantic herring. *ICES Journal of Marine Science*, 72(6), 1790-1801.

Bektas, Y., and Belduz, A.O. (2008). Molecular phylogeny of Turkish *Trachurus* species (*Perciformes: Caran-*

gidae) inferred from mitochondrial DNA analyses. *Journal of Fish Biology*, 73(5), 1228-1248.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. (2004). GENETIX 4.05, Windows TM software for population genetics. Laboratoire génome, populations, interactions, CNRS UMR, 5000.

Benny, P.N. (2002). Variability of Western Indian Ocean currents. *Western Indian Ocean Journal of Marine Science,* 1, 81-90.

Berg, H., Francis, J., Souter, P. (2002). Support to marine research for sustainable management of marine and coastal resources in the Western Indian Ocean. *Ambio*, 31, 597-601.

Bermingham, E., and Moritz, C. (1998). Comparative phylogeography: concepts and applications. *Molecular Ecology*, 7(4), 367-369.

Bermingham, E., McCafferty, S.S., and Martin, A.P. (1997). Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. *Molecular systematics of fishes*.

Bernardon, M., and Vall, M.O.M. (2004). *The mullet in Mauritania: biology, ecology, fisheries and management* (p.53). FIBA, Banc d'Arguin International Foundation.

Bernatchez, L. (2001). The evolutionary history of brown trout (Salmo trutta L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution*, 55(2), 351-379.

Bernatchez, L., Glémet, H., Wilson, C.C., and Danzmann, R.G. (1995). Introgression and fixation of Arctic char (Salvelinus alpinus) mitochondrial genome in an allopatric population of brook trout (Salvelinus fontinalis). *Canadian Journal of Fisheries and Aquatic Sciences*, 52(1), 179-185.

Berry, O., England, P., Marriott, R.J., Burridge, C.P., and Newman, S.J. (2012). Understanding age-specific dispersal in fishes through hydrodynamic modelling, genetic simulations and microsatellite DNA analysis. *Molecular Ecology*, 21: 2145–2159.

Bhat, S., Amundsen, P.A., Knudsen, R., Gjelland, K.Ø., Fevolden, S.E., Bernatchez, L., and Præbel, K. (2014). Speciation reversal in European whitefish (Coregonus lavaretus (L.)) caused by competitor invasion. *PLoS One*, 9(3), e91208.

Bianchi, G. (1992a). Study of the demersal assemblages of the continental shelf and upper slope off Congo and Gabon, based on the trawl surveys of the RV'Dr Fridtjof Nansen'.

Bianchi, G. (1992b). Demersal assemblages of the continental shelf and upper slope of Angola. *Marine Ecology Progress Series,* 81: 101–120.

Bierne, N., Welch, J., Loire, E., Bonhomme, F., and David, P. (2011). The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, 20(10), 2044-2072.

Billups, K., Ravelo, A.C., Zachos, J.C., and Norris, R.D. (1999). Link between oceanic heat transport, thermohaline circulation, and the Intertropical Convergence Zone in the early Pliocene Atlantic. *Geology*, 27(4), 319-322.

Bird, C.E., Holland, B.S., Bowen, B.W., and Toonen, R.J. (2007). Contrasting phylogeography in three endemic Hawaiian limpets (Cellana spp.) with similar life histories. *Molecular Ecology*, 16, 3173-3186.

Bird, C.E., Holland, B.S., Bowen, B.W., and Toonen, R.J. (2011). Diversification of sympatric broadcast-spawning limpets (Cellana spp.) within the Hawaiian archipelago. *Molecular Ecology,* 20, 2128-2141.

Blaber, S.J.M. (1977). The feeding ecology and relative abundance of mullet (Mugilidae) in Natal and Pondoland estuaries. *Biological Journal of the Linnean Society*, 9(3), 259-275.

Blel, H., Chatti, N., Besbes, R., Farjallah, S., Elouaer, A., Guerbej, H., and Said, K. (2008). Phylogenetic relationships in grey mullets (Mugilidae) in a Tunisian lagoon. *Aquaculture research*, 39(3), 268-275.

Boateng, M. (2019). Fishes of Ghana. Version 1.3. Occurrence dataset. In *Department of Marine and Fisheries Sciences, University of Ghana*. Available at: https://doi.org/10.15468/pgesnw (last accessed on 12 April 2019 via GBIF.org).

Bohonak, A.J. (1999). Dispersal, gene flow, and population structure. *Quarterly review of biology*, 21-45.

Boisvert, W.E. (1967). *Major currents in the North and South Atlantic Oceans between 64 N and 60 S*. Naval Oceanographic Office.

Bonnet, T., Leblois, R., Rousset, F., and Crochet, P.A. (2017). A reassessment of explanations for discordant introgressions of mitochondrial and nuclear genomes. *Evolution*, 71(9), 2140-2158.

Borrero-Pérez, G.H., González-Wangüemert, M., Marcos, C., and Pérez-Ruzafa, A. (2011). Phylogeography of the Atlanto-Mediterranean sea cucumber Holothuria (Holothuria) mammata: the combined effects of historical processes and current oceanographical pattern. *Molecular ecology*, 20(9), 1964-1975.

Borsa, P. (2002). Allozyme, mitochondrial-DNA, and morphometric variability indicate cryptic species of anchovy (Engraulis encrasicolus). *Biological Journal of the Linnean Society*, 75(2): 261–269.

Borsa, P., Hsiao, D.R., Carpenter, K.E., and Chen, W. J. (2013). Cranial morphometrics and mitochondrial DNA sequences distinguish cryptic species of the longface emperor (Lethrinus olivaceus), an emblematic fish of Indo-West Pacific coral reefs. *Comptes rendus biologies*, 336(10), 505-514.

Borsa, P., Naciri, M., Bahri, L., Chikhi, L., Garcia de Leon, F.J., Kotoulas, G., and Bonhomme, F. (1997). Atlanto-Mediterranean species (fishes and invertebrates). *Life and Environment (France).*

Bourjea, J., Lapegue, S., Gagnevin, L., Broderick, D., Mortimer, J.A., Ciccione, S., ... and Grizel, H. (2007). Phylogeography of the green turtle, Chelonia mydas, in the Southwest Indian Ocean. *Molecular Ecology*, 16(1), 175-186.

Bourret, V., Kent, M.P., Primmer, C.R., Vasemägi, A., Karlsson, S., Hindar, K., ... and Lien, S. (2013). SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of A tlantic salmon (S almo salar). *Molecular ecology*, 22(3), 532-551.

Bowen, B.W., Gaither, M.R., DiBattista, J.D., Iacchei, M., Andrews, K.R., Grant, W.S., ... and Briggs, J.C. (2016). Comparative phylogeography of the ocean planet. *Proceedings of the National Academy of Sciences*, 113(29), 7962-7969.Boyd, A.J., Salat, J., and Masó, M. (1987). The seasonal intrusion of relatively saline water on the shelf off northern and central Namibia. *South African Journal of Marine Science*, 5 (1), 107-120.

Bowen, B.W., Gaither, M.R., DiBattista, J.D., Iacchei, M., Andrews, K.R., Grant, W.S., ... and Briggs, J.C. (2016). Comparative phylogeography of the ocean planet. *Proceedings of the National Academy of Sciences*, 113(29), 7962-7969.

Boyer, D., Cole, J., and Bartholomae, C. (2000). Southwestern Africa: northern Benguela current region: an environmental evaluation. *Marine Pollution Bulletin*, 41, 123–140.

Boyle, P.R. (ed.). (2000). Cephalopod resource dynamics: patterns in environmental and genetic variation, 1997–2000. Final Report for EU Integrated Project FAIR-CT96-1520. Aberdeen: University of Aberdeen, UK.

Brawand, D., Wagner, C.E., Li, Y.I., Malinsky, M., Keller, I., Fan, S., ... and Turner-Maier, J. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature*, 513(7518), 375.

Bremer, J.R.A., Viñas, J., Mejuto, J., Ely, B., and Pla, C. (2005). Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. *Molecular phylogenetics and evolution*, 36(1), 169-187.

Briggs, J.C. (1960). Fishes of worldwide (circumtropical) distribution. *Copeia*, 1960, 171-180

Briggs, J.C. (1974) Marine zoogeography. NewYork: McGraw-Hill

Briggs, J.C. (1987a). Antitropical distribution and evolution in the Indo-West Pacific Ocean. *Systematic Zoology*, 36(3), 237-247.

Briggs, J.C. (1987b). Antitropicality and vicariance. *Systematic Zoology*, 36(2), 206-207.

Briggs, J.C. (1995). *Global biogeography* (Vol. 14). Elsevier.

Briggs, J.C., and Bowen, B.W. (2012). A realignment of marine biogeographic provinces with particular reference to fish distributions. *Journal of Biogeography*, 39(1), 12-30.

Brill, R.W., and Lutcavage, M.E. (2001). Understanding environmental influences on movements and depth distributions of tunas and billfishes can significantly improve population assessments. In *American Fisheries Society Symposium* (pp. 179-198). American Fisheries Society.

Bruch, A.A., Uhl, D., and Mosbrugger, V. (2007). Miocene climate in Europe—patterns and evolution: a first synthesis of NECLIME.

Brulhet, J. (1975). Observations on the biology of Mugil cephalus ashentensis and the possibility of its aquaculture on the Mauritanian coast. *Aquaculture*, 5(3), 271-281.

Brusle, J. (1981). *Sexuality and biology of reproduction in grey mullets*. In: O.H. Oren (Ed.), Aquaculture of Grey Mullets, Cambridge University Press, Cambridge (1981), pp. 99-154

Burford, M.O., Carr, M.H., and Bernardi, G. (2011). Age-structured genetic analysis reveals temporal and geographic variation within and between two cryptic rockfish species. *Marine Ecology Progress Series*, 442, 201-215.

Burridge, C.P. (2002). Antitropicality of Pacific fishes: molecular insights. *Environmental Biology of Fishes*, 65

(2), 151-164.

Burridge, C.P., Hurt, A.C., Farrington, L.W., Coutin, P.C., and Austin, C.M. (2004). Stepping stone gene flow in an estuarine-dwelling sparid from south-east Australia. *Journal of Fish Biology*, 64(4), 805-819.

Burrows, M.T., Schoeman, D.S., Buckley, L.B., Moore, P., Poloczanska, E.S., Brander, K.M., ... and Holding, J. (2011). The pace of shifting climate in marine and terrestrial ecosystems. *Science*, 334(6056), 652-655.

Burrows, M.T., Schoeman, D.S., Richardson, A.J., Molinos, J.G., Hoffmann, A., Buckley, L.B., ... and Halpern, B.S. (2014). Geographical limits to species-range shifts are suggested by climate velocity. *Nature*, 507 (7493), 492.

Butlin, R.K., Galindo, J., and Grahame, J.W. (2008). Sympatric, parapatric or allopatric: the most important way to classify speciation?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363 (1506), 2997-3007.

Caddy, J.F., and Rodhouse, P.G. (1998). Cephalopod and groundfish landings: evidence for ecological change in global fisheries?. *Reviews in Fish Biology and Fisheries*, 8(4), 431-444.

Caldara, F., Bargelloni, L., Ostellari, L., Penzo, E., Colombo, L., and Patarnello, T. (1996). Molecular phylogeny of grey mullets based on mitochondrial DNA sequence analysis: evidence of a differential rate of evolution at the intrafamily level. *Molecular Phylogenetics and Evolution*, 6(3), 416-424.

Calderón, I., Giribet, G., and Turon, X. (2008). Two markers and one history: phylogeography of the edible common sea urchin Paracentrotus lividus in the Lusitanian region. *Marine Biology*, 154(1), 137-151.

Campo, D., Molares, J., Garcia, L., Fernandez-Rueda, P., Garcia-Gonzalez, C., and Garcia-Vazquez, E. (2010). Phylogeography of the European stalked barnacle (Pollicipes pollicipes): identification of glacial refugia. *Marine Biology*, 157(1), 147.

Canales-Aguirre, C.B., Ferrada, S., and Galleguillos, R. (2010). Isolation and characterization of microsatellite loci for the jack mackerel (*Trachurus murphyi* Nichols, 1920). *Conservation genetics*, 11(3), 1235-1237.

Candela, J. (2001). .7 Mediterranean water and global circulation. In *International Geophysics* (Vol. 77, pp. 419-XLVIII). Academic Press.

Cardenas, L., Hernandez, C.E., Poulin, E., Magoulas, A., Kornfield, I., and Ojeda, F.P. (2005). Origin, diversification, and historical biogeography of the genus *Trachurus (Perciformes: Carangidae)*. *Molecular Phylogenetics and Evolution*, 35(2), 496-507.

Carpenter, K.E., and Allen, G.R. (1989). FAO species catalogue (Vol. 9). Food and Agriculture Organisation, Rome.

Carpenter, K.E., & Johnson, G.D. (2002). A phylogeny of sparoid fishes (Perciformes, Percoidei) based on morphology. *Ichthyological Research*, 49(2), 114-127.

Carpenter, K.E., & Russell, B. (2014). Dentex dentex. In *Species*. IUCN.

Carreras-Carbonell, J., Macpherson, E., and Pascual, M. (2006). Population structure within and between subspecies of the Mediterranean triplefin fish Tripterygion delaisi revealed by highly polymorphic microsatellite loci. *Molecular Ecology*, 15, 3527-3529.

Carvalho, G.R., & Hauser, L. (1995). Molecular genetics and the stock concept in fisheries. In *Molecular genetics in fisheries* (pp. 55-79). Springer, Dordrecht.

Castiglia, R. (2014). Sympatric sister species in rodents are more chromosomally differentiated than allopatric ones: implications for the role of chromosomal rearrangements in speciation. *Mammal Review*, 44(1), 1-4.

Catanach, A., Crowhurst, R., Deng, C., David, C., Bernatchez, L., and Wellenreuther, M. (2019). The genomic pool of standing structural variation outnumbers single nucleotide polymorphism by three-fold in the marine teleost Chrysophrys auratus. *Molecular ecology*. 28(6), 1210-1223

Catarino, D., Knutsen, H., Veríssimo, A., Olsen, E.M., Jorde, P.E., Menezes, G., ... and Danovaro, R. (2015). The Pillars of Hercules as a bathymetric barrier to gene flow promoting isolation in a global deep-sea shark (C entroscymnus coelolepis). *Molecular ecology*, 24(24), 6061-6079.

Cataudella, S., Civitelli, M.V., and Capanna, E. (1974). Chromosome complements of the Mediterranean mullets (Pisces Perciformes). *Caryologia*, 27(1), 93-105.

Cavalli-Sforza, L.L., and Edwards, A.W. (1967). Phylogenetic analysis: models and estimation procedures. *Evolution*, 21(3), 550-570.

Chabanet, P., and Durville, P. (2005). Reef fish inventory of Juan de Nova's natural park (western Indian Ocean). *Western Indian Ocean Journal of Marine Science*, 4(2), 145-162.

Chakraborty, R. (1992) Sample-size requirements for addressing the population genetic issues of forensic use of dna typing. *Human Biology*, 64, 141-159.

Chang C.W., Tzeng, W.N., and Lee, Y.C., (2000) Recruitment and hatching dates of grey mullet (Mugil cephalus L.) juveniles in the Tanshui estuary of northwest Taiwan. *Zoological Studies* 39, no. 2: 99-106.

Chang, C.W., and Iizuka, Y. (2012). Estuarine use and movement patterns of seven sympatric Mugilidae fishes: the Tatu Creek estuary, central western Taiwan. *Estuarine, Coastal and Shelf Science*, 106, 121-126.

Chapuis, M.P., and Estoup, A. (2006). Microsatellite null alleles and estimation of population differentiation. *Molecular biology and evolution*, 24(3), 621-631.

Charlesworth, B., and Barton, N.H. (2018). The spread of an inversion with migration and selection. *Genetics*, 208(1), 377-382.

Charrier, G., Chenel, T., Durand, J.D., Girard, M., Quiniou, L., and Laroche, J. (2006). Discrepancies in phylogeographical patterns of two European anglerfishes (Lophius budegassa and Lophius piscatorius). *Molecular phylogenetics and evolution*, 38(3), 742-754.

Chen, W.J., Ruiz-Carus, R., and Ortí, G. (2007). Relationships among four genera of mojarras (Teleostei: Perciformes: Gerreidae) from the western Atlantic and their tentative placement among percomorph fishes. *Journal of Fish Biology*, 70, 202-218.

Chenoweth, S.F., Hughes, J.M., Keenan, C.P., and Lavery, S. (1998a). When oceans meet: a teleost shows secondary intergradation at an Indian–Pacific interface. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1394), 415-420.

Chenoweth, S.F., Hughes, J.M., Keenan, C.P., and Lavery, S. (1998b). Concordance between dispersal and mitochondrial gene flow: isolation by distance in a tropical teleost, Lates calcarifer (Australian barramundi). *Heredity*, 80(2), 187.

Chevin, L.M., Lande, R., and Mace, G.M. (2010). Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS biology*, 8(4), e1000357.

Chevolot, M., Hoarau, G., Rijnsdorp, A.D., Stam, W.T., and Olsen, J.L. (2006). Phylogeography and population structure of thornback rays (Raja clavata L., Rajidae). *Molecular Ecology*, 15(12), 3693-3705.

Chiang, H.C., Hsu, C.C., Wu, G.C.C., Chang, S.K., and Yang, H.Y. (2008). Population structure of bigeye tuna (Thunnus obesus) in the Indian Ocean inferred from mitochondrial DNA. *Fisheries Research,* 90, 305-312.

Chiang, J.C., Kushnir, Y., and Giannini, A. (2002). Deconstructing Atlantic Intertropical Convergence Zone variability: Influence of the local cross-equatorial sea surface temperature gradient and remote forcing from the eastern equatorial Pacific. *Journal of Geophysical Research: Atmospheres*, 107(D1), ACL-3.

Chiba, S.N., Iwatsuki, Y., Yoshino, T., & Hanzawa, N. (2009). Comprehensive phylogeny of the family Sparidae (Perciformes: Teleostei) inferred from mitochondrial gene analyses. *Genes & genetic systems*, 84(2), 153-170.

Chikhi, L., Bonhomme, F., and Agnèse, J.F. (1998). Low genetic variability in a widely distributed and abundant clupeid species, Sardinella aurita. New empirical results and interpretations. *Journal of Fish Biology*, 52(5), 861-878.

Chlaida, M., Kifani, S., Lenfant, P., and Ouragh, L. (2006). First approach for the identification of sardine populations Sardina pilchardus (Walbaum 1792) in the Moroccan Atlantic by allozymes. *Marine Biology*, 149(2), 169.

Christie, M.R., Johnson, D.W., Stallings, C.D., and Hixon, M.A. (2010). Self-recruitment and sweepstakes reproduction amid extensive gene flow in a coral-reef fish. *Molecular Ecology,* 19, 1042-1057.

Chu, K.H., Li, C.P., Tam, Y.K., and Lavery, S. (2003). Application of mitochondrial control region in population genetic studies of the shrimp Penaeus. *Molecular Ecology Notes*, *3*(1), 120-122.

Chubb, A.L., Zink, R.M., and Fitzsimons, J.M. (1998). Patterns of mtDNA variation in Hawaiian freshwater fishes: the phylogeographic consequences of amphidromy. *Journal of Heredity*, 89(1), 8-16.

Cimmaruta, R., Bondanelli, P., and Nascetti, G. (2005). Genetic structure and environmental heterogeneity in the European hake (*Merluccius merluccius*). *Molecular ecology*, 14(8), 2577-2591.

Cimmaruta, R., Bondanelli, P., Ruggi, A., and Nascetti, G. (2008). Genetic structure and temporal stability in the horse mackerel (*Trachurus trachurus*). *Fisheries Research*, 89(2), 114-121.

Clauzon, G., Suc, J.P., Gautier, F., Berger, A., and Loutre, M.F. (1996). Alternate interpretation of the Messini-

an salinity crisis: Controversy resolved?. *Geology*, 24(4), 363-366.

Coleman, F.C., and Williams, S.L. (2002). Overexploiting marine ecosystem engineers: potential consequences for biodiversity. *Trends in Ecology and Evolution*, 17(1), 40-44.

Colleoni, F., Masina, S., Negri, A., and Marzocchi, A. (2012). Plio–Pleistocene high–low latitude climate interplay: A Mediterranean point of view. *Earth and planetary science letters*, 319, 35-44.F

Collins, R.A., and Cruickshank, R.H. (2013). The seven deadly sins of DNA barcoding. *Molecular Ecology Resources*, 13(6): 969–975.

Combosch, D.J., Lemer, S., Ward, P.D., Landman, N.H., and Giribet, G. (2017). Genomic signatures of evolution in Nautilus—an endangered living fossil. *Molecular ecology*, 26(21), 5923-5938.

Comesaña, A.S., Martínez-Areal, M.T., and Sanjuan, A. (2008). Genetic variation in the mitochondrial DNA control region among horse mackerel (*Trachurus trach*urus) from the Atlantic and Mediterranean areas. *Fisheries Research*, 89(2), 122-131.

Corander, J., Waldmann, P., Marttinen, P., and Sillanpaa, M.J. (2004). BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics,* 20, 2363-2369.

Coscia, I., Robins, P.E., Porter, J.S., Malham, S.K., and Ironside, J.E. (2013). Modelled larval dispersal and measured gene flow: seascape genetics of the common cockle *Cerastoderma edule* in the southern Irish Sea. *Conservation Genetics*, 14(2), 451-466.

Cowen, R.K., Lwiza, K.M., Sponaugle, S., Paris, C.B., and Olson, D.B. (2000). Connectivity of marine populations: open or closed?. *Science*,287(5454), 857-859.

Coyne, J.A., and Orr, H.A. (1997). "Patterns of speciation in Drosophila" revisited. *Evolution*, 51(1), 295-303.

Coyne, J.A., and Orr, H.A. (2004). Speciation. Sinauer. *Sunderland, MA*.

Craig, M.T., and Hastings, P.A. (2007). A molecular phylogeny of the groupers of the subfamily Epinephelinae (Serranidae) with a revised classification of the Epinephelini. *Ichthyological Research*, 54(1), 1-17.

Craig, M.T., Eble, J.A., Bowen, B.W., and Robertson, D. (2007). High genetic connectivity across the Indian and Pacific Oceans in the reef fish Myripristis berndti (Holocentridae). *Marine Ecology Progress Series*, 334(3), 245-254.

Crosetti, D. (2016). Current State of Grey Mullet Fisheries and Culture. In.: Crosetti, D., Blaber, S.(Eds.), Biology, Ecology and Culture of Grey Mullet (Mugilidae).

Crosetti, D., Avise, J.C., Placidi, F., Rossi, A.R., and Sola, L. (1993). Geographic variability in the grey mullet Mugil cephalus: preliminary results of mtDNA and chromosome analyses. In *Genetics in Aquaculture* (pp. 95-101).

Crosetti, D., Nelson, W.S., and Avise, J.C. (1994). Pronounced genetic structure of mitochondrial DNA among populations of the circumglobally distributed grey mullet (Mugil cephalus). *Journal of fish biology*, 44 (1), 47-58.

Crozier, R.H., & Crozier, Y.C. (1993). The mitochondrial genome of the honeybee Apis mellifera: complete sequence and genome organization. *Genetics*, 133(1), 97-117.

Cunningham, C.W., and Collins, T.M. (1998). Beyond area relationships: extinction and recolonization in molecular marine biogeography. In *Molecular approaches to ecology and evolution* (pp. 297-321). Birkhäuser, Basel.

D'Amato, M.E., Harkins, G.W., de Oliveira, T., Teske, P.R., and Gibbons, M.J. (2008). Molecular dating and biogeography of the neritic krill Nyctiphanes. *Marine Biology*, 155(2), 243-247.

Dannewitz, J., Maes, G.E., Johansson, L., Wickström, H., Volckaert, F.A., & Järvi, T. (2005). Panmixia in the European eel: a matter of time…. *Proceedings of the Royal Society B: Biological Sciences*, 272(1568), 1129-1137.

Daemen, E., Cross, T., Ollevier, F., and Volckaert, F. (2001). Analysis of the genetic structure of European eel (Anguilla anguilla) using microsatellite DNA and mtDNA markers. *Marine Biology*, 139(4), 755-764.

Dai, L., Zheng, X., Kong, L., and Li, Q. (2012). DNA barcoding analysis of Coleoidea (Mollusca: Cephalopoda) from Chinese waters. *Molecular Ecology Resources*, 12: 437–447.

Davis, M.B., Shaw, R.G., and Etterson, J.R. (2005). Evolutionary responses to changing climate. *Ecology*, 86(7), 1704-1714.

Dawson, M.N. (2012). Parallel phylogeographic structure in ecologically similar sympatric sister taxa. *Molecular Ecology*, 21(4), 987-1004.

De Beer, C.L. (2014). *Genetic and morphometric variation of Octopus vulgaris in the Benguela Current Region* (Doctoral dissertation, Rhodes University).

De Heij, A., and Baayen, R.P. (2005). Seasonal distribution of cephalopod species living in the central and southern North Sea. *Basteria*, 69: 91–119.

De Leeuw, A., Mandic, O., de Bruijn, H., Marković, Z., Reumer, J., Wessels, W., ... and Krijgsman, W. (2011). Magnetostratigraphy and small mammals of the Late Oligocene Banovići basin in NE Bosnia and Herzegovina. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 310(3-4), 400-412.

De Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation. *Endless forms: species and speciation*.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic biology*, 56(6), 879-886.

De Wit, P., and Palumbi, S.R. (2013). Transcriptome-wide polymorphisms of red abalone (H aliotis rufescens) reveal patterns of gene flow and local adaptation. *Molecular Ecology*, 22(11), 2884-2897.

Dean, W.E., Hay, W.W., and Sibuet, J.C. (1984). Geologic evolution, sedimentation and paleoenvirouments of the Angola Basin and adjacent Walvis Ridge: synthesis of results of Deep Sea Drilling Project Leg 75. *Initial Reports of the Deep Sea Drilling Project*, *75*(JUN), 509-544.

Delgado, J.V., Molina, A., Lobillo, J., Alonso, A., and Camacho, M.E. (1992). Morphometrical study on the chromosomes of three species of mullet (Teleostei, Mugilidae). *Caryologia*, 45(3-4), 263-271.

Delport, W., Poon, A.F., Frost, S.D., and Kosakovsky Pond, S.L. (2010). Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*, 26(19), 2455-2457.

Díaz-Jaimes, P., Uribe-Alcocer, M., Rocha-Olivares, A., García-de-León, F.J., Nortmoon, P., and Durand, J.D. (2010). Global phylogeography of the dolphinfish (Coryphaena hippurus): the influence of large effective population size and recent dispersal on the divergence of a marine pelagic cosmopolitan species. *Molecular Phylogenetics and Evolution*, 57(3), 1209-1218.

Dieckmann, U., Doebeli, M., Metz, J.A., and Tautz, D. (Eds.). (2004). *Adaptive speciation*. Cambridge University Press.

Diester-Haass, L., Meyers, P.A., and Rothe, P. (1990). Miocene history of the Benguela Current and Antarctic ice volumes: Evidence from rhythmic sedimentation and current growth across the Walvis Ridge (Deep Sea Drilling Project Sites 362 and 532). *Paleoceanography and Paleoclimatology*, 5(5), 685-707.

Diester-Haass, L., Meyers, P.A., and Rothe, P. (1992). The Benguela Current and associated upwelling on the southwest African Margin: a synthesis of the Neogene-Quaternary sedimentary record at DSDP sites 362 and 532. *Geological Society, London, Special Publications*, 64(1), 331-342.

Diester-Haass, L., Meyers, P.A., and Vidal, L. (2002). The late Miocene onset of high productivity in the Benguela Current upwelling system as part of a global pattern. *Marine Geology*, 180(1-4), 87-103.

Dillane, E., Galvin, P., Coughlan, J., Lipinski, M., and Cross, T.F. (2005). Genetic variation in the lesser flying squid Todaropsis eblanae (Cephalopoda, Ommastrephidae) in east Atlantic and Mediterranean waters. *Marine Ecology Progress Series*, 292, 225-232.

Diniz, F.M., Maclean, N., Ogawa, M., Cintra, I.H., and Bentzen, P. (2005). The hypervariable domain of the mitochondrial control region in Atlantic spiny lobsters and its potential as a marker for investigating phylogeographic structuring. *Marine Biotechnology*, 7(5), 462-473.

Djakouré, S., Penven, P., Bourlès, B., Koné, V., and Veitch, J. (2017). Respective Roles of the Guinea Current and Local Winds on the Coastal Upwelling in the Northern Gulf of Guinea. *Journal of Physical Oceanography*, 47(6), 1367-1387.

Dobzhansky, T.H. (1936). Studies on hybrid sterility. II. Localization of sterility factors in Drosophila pseudoobscura hybrids. *Genetics*, 21(2), 113.

Doiron, S., Bernatchez, L., and Blier, P.U. (2002). A comparative mitogenomic analysis of the potential adaptive value of arctic charr mtDNA introgression in brook charr populations (Salvelinus fontinalis Mitchill). *Molecular biology and evolution*, 19(11), 1902-1909.

Domingues, V.S., Bucciarelli, G., Almada, V.C., and Bernardi, G. (2005). Historical colonization and demography of the Mediterranean damselfish, Chromis chromis. *Molecular Ecology*, 14(13), 4051-4063.

Domingues, V.S., Faria, C., Stefanni, S., Santos, R.S., Brito, A., and Almada, V.C. (2007b). Genetic divergence in the Atlantic–Mediterranean Montagu's blenny, Coryphoblennius galerita (Linnaeus 1758) revealed by molecular and morphological characters. *Molecular Ecology*, 16(17), 3592-3605.

Domingues, V.S., Santos, R.S., Brito, A., Alexandrou, M., and Almada, V.C. (2007a). Mitochondrial and nuclear markers reveal isolation by distance and effects of Pleistocene glaciations in the northeastern Atlantic

and Mediterranean populations of the white seabream (Diplodus sargus, L.). *Journal of Experimental Marine Biology and Ecology*, 346(1-2), 102-113.

Donoghue, P.C., and Yang, Z. (2016). The evolution of methods for establishing evolutionary time-scales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1699), 20160020.

Doubleday, Z.A., Prowse, T.A., Arkhipkin, A., Pierce, G.J., Semmens, J., Steer, M., ... and Gillanders, B. M. (2016). Global proliferation of cephalopods. *Current Biology*, 26(10), R406-R407.

Dowsett, H., and Willard, D. (1996). Southeast Atlantic marine and terrestrial response to middle Pliocene climate change. *Marine Micropaleontology*, 27(1-4), 181-193.

Dowsett, H., Barron, J., and Poore, R. (1996). Middle Pliocene sea surface temperatures: a global reconstruction. *Marine Micropaleontology*, 27(1-4), 13-25.

Dowsett, H.J., Cronin, T.M., Poore, R.Z., Thompson, R.S., Whatley, R.C., and Wood, A.M. (1992). Micropaleontological evidence for increased meridional heat transport in the North Atlantic Ocean during the Pliocene. *Science*, 258(5085), 1133-1135.

Duarte, A., Fielding, P., Sowman, M., Bergh, M. (2005). Overview and analysis of socio-economic and fisheries information to promote the management of artisanal fisheries in the BCLME region-Angola. Unpublished Final Report No. LMR/AFSE/03/01/B. *Cape Town.*

Duchêne, S., Lanfear, R., and Ho, S.Y. (2014). The impact of calibration and clock-model choice on molecular estimates of divergence times. *Molecular phylogenetics and evolution*, 78, 277-289.

Duda, T.F., and Palumbi, S.R. (1999). Population structure of the black tiger prawn, Penaeus monodon, among western Indian Ocean and western Pacific populations. *Marine Biology,* 134, 705-710.

Dunn, M.R. (1999). Aspects of the stock dynamics and exploitation of cuttlefish, Sepia officinalis (Linnaeus, 1758), in the English Channel. *Fisheries Research*, 40(3), 277-293.

Duran, S., Palacin, C., Becerro, M.A., Turon, X., and Giribet, G. (2004). Genetic diversity and population structure of the commercially harvested sea urchin Paracentrotus lividus (Echinodermata, Echinoidea). *Molecular ecology*, 13(11), 3317-3328.

Durand, J. D., and Borsa, P. (2015). Mitochondrial phylogeny of grey mullets (Acanthopterygii: Mugilidae) suggests high proportion of cryptic species. *Comptes Rendus Biologies*, *338*(4), 266-277.

Durand, J. D., and Whitfield, A. K. (2015). Biogeography and Distribution of Mugilidae in the Western, Central and Southern Regions of Africa. In *Biology, Ecology and Culture of Grey Mullets (Mugilidae)* (pp. 111-124). CRC Press.

Durand, J.D., Blel, H., Shen, K.N., Koutrakis, E.T., and Guinand, B. (2013). Population genetic structure of Mugil cephalus in the Mediterranean and Black Seas: a single mitochondrial clade and many nuclear barriers. *Marine Ecology Progress Series*, 474, 243-261.

Durand, J.D., Chen, W.J., Shen, K.N., Fu, C., and Borsa, P. (2012b). Genus-level taxonomic changes implied by the mitochondrial phylogeny of grey mullets (Teleostei: Mugilidae). *Comptes rendus biologies*, 335(10-11), 687-697.

Durand, J.D., Guinand, B., Dodson, J.J., and Lecomte, F. (2013). Pelagic life and depth: coastal physical features in West Africa shape the genetic structure of the Bonga Shad, Ethmalosa fimbriata. *PloS one*, 8 (10), e77483.

Durand, J.D., Hubert, N., Shen, K.N., and Borsa, P. (2017). DNA barcoding grey mullets. *Reviews in Fish Biology and Fisheries*, 27(1), 233-243.

Durand, J.D., Shen, K.N., Chen, W.J., Jamandre, B.W., Blel, H., Diop, K., ... and Borsa, P. (2012a). Systematics of the grey mullets (Teleostei: Mugiliformes: Mugilidae): molecular phylogenetic evidence challenges two centuries of morphology-based taxonomy. *Molecular Phylogenetics and Evolution*, 64(1), 73-92.

Durand, J.D., Tine, M., Panfili, J., Thiaw, O.T., and Lae, R. (2005b). Impact of glaciations and geographic distance on the genetic structure of a tropical estuarine fish, Ethmalosa fimbriata (Clupeidae, S. Bowdich, 1825). *Molecular Phylogenetics and Evolution*, 36(2), 277-287.

Ebisawa, A. (2006). Reproductive and sexual characteristics in five Lethrinus species in waters off the Ryukyu Islands. *Ichthyological Research,* 53, 269-280.

Ebisawa, A., and Ozawa, T. (2009). Life-history traits of eight Lethrinus species from two local populations in waters off the Ryukyu Islands. *Fisheries Science*, 75, 553-566.

Economidis, P.S., Dimitriou, E., Pagoni, R., Michaloudi, E., and Natsis, L. (2000). Introduced and translocated fish species in the inland waters of Greece. *Fisheries Management and Ecology*, 7(3), 239-250.

Ekman, S. (1953). *Zoogeography of the sea* (Vol. 9). Sidgwick and Jackson.

Elderfield, H., Ferretti, P., Greaves, M., Crowhurst, S., McCave, I.N., Hodell, D.A., and Piotrowski, A.M. (2012). Evolution of ocean temperature and ice volume through the mid-Pleistocene climate transition. *Science*, 337(6095), 704-709.

El-Geziry, T.M., and Bryden, I.G. (2010). The circulation pattern in the Mediterranean Sea: issues for modeller consideration. *Journal of Operational Oceanography*, 3(2), 39-46.

Ellegren, H., Smith, N.G., and Webster, M.T. (2003). Mutation rate variation in the mammalian genome. *Current opinion in genetics and development*, 13(6), 562-568.

El-Sayed, A.F.M. (2015). Regional review on status and trends in aquaculture development in the near east and North Africa. *FAO Fisheries and Aquaculture Circular*, (C1135/6), I.

Ely, B., Stoner, D.S., Bremer, A.J., Dean, J.M., Addis, P., Cau, A., ... and Scott, K. (2002). Analyses of nuclear ldhA gene and mtDNA control region sequences of Atlantic northern bluefin tuna populations. *Marine Biotechnology*, 4(6), 583-588.

Emery, K.O., and Uchupi, E. (1984). *The geology of the Atlantic Ocean*. Springer Science and Business Media.

Endler, J.A. (1977). *Geographic variation, speciation, and clines* (No. 10). Princeton University Press.

Endo, T., Ikeo, K., and Gojobori, T. (1996). Large-scale search for genes on which positive selection may operate. *Molecular biology and evolution*, 13(5), 685-690.

Engels, B. (2016). HWxtest: exact tests for Hardy-Weinberg proportions. *R package version*, *1*(7).

Erguden, D., Gurlek, M., Yaglioglu, D., and Turan, C. (2010). Genetic Identification and Taxonomic Relationship of Mediterranean Mugilid Species Based on Mitochondrial 168 rDNA Sequence Data. *Journal of Animal and Veterinary Advances*, 9(2), 336-341.

Eschmeyer W.N., and Fong, J.D. (2015). Catalog of Fishes. Available on line at: http://research.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp. (Last accessed 20/11/18)

Estoup, A., Jarne, P., and Cornuet, J.M. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular ecology*, 11(9), 1591-1604.

Evans, B.S., Sweijd, N.A., Bowie, R.C.K., Cook, P.A., and Elliott, N.G. (2004). Population genetic structure of the perlemoen Haliotis midae in South Africa: evidence of range expansion and founder events. *Marine Ecology Progress Series*, 270, 163-172.

Excoffier, L., and Lischer, H.E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources,* 10(3), 564-567.

Excoffier, L., Smouse, P.E., and Quattro, J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131(2): 479–491.

Fairbanks, R.G. (1989). A 17,000-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature*, 342(6250), 637.

Falk, T.M., Teugels, G.G., Abban, E.K., Villwock, W., and Renwrantz, L. (2003). Phylogeographic patterns in populations of the black-chinned tilapia complex (Teleostei, Cichlidae) from coastal areas in West Africa: support for the refuge zone theory. *Molecular Phylogenetics and Evolution*, 27(1), 81-92.

Fan, S., and Meyer, A. (2014). Evolution of genomic structural variation and genomic architecture in the adaptive radiations of African cichlid fishes. *Frontiers in genetics*, 5, 163.

Faria, R., Chaube, P., Morales, H.E., Larsson, T., Lemmon, A.R., Lemmon, E.M., ... and Westram, A.M. (2019). Multiple chromosomal rearrangements in a hybrid zone between Littorina saxatilis ecotypes. *Molecular ecology*. 28(6), 1375-1393.

Faria, R., Weiss, S., and Alexandrino, P. (2012). Comparative phylogeography and demographic history of European shads (Alosa alosa and A. fallax) inferred from mitochondrial DNA. *BMC evolutionary biology*, 12(1), 194.

Fauvelot, C., Bernardi, G., and Planes, S. (2003). Reductions in the mitochondrial DNA diversity of coral reef fish provide evidence of population bottlenecks resulting from Holocene sea-level change. *Evolution*, 57(7), 1571-1583.

Feder, J.L., Berlocher, S.H., Roethele, J.B., Dambroski, H., Smith, J.J., Perry, W.L., ... and Aluja, M. (2003). Allopatric genetic origins for sympatric host-plant shifts and race formation in Rhagoletis. *Proceedings of the National Academy of Sciences*, 100(18), 10314-10319.

Fennel, W. (1999). Theory of the Benguela upwelling system. *Journal of Physical Oceanography*, 29(2), 177-190.

Féral, J.P. (2002). How useful are the genetic markers in attempts to understand and manage marine biodiversity?. *Journal of Experimental Marine Biology and Ecology*, 268(2), 121-145.

Ferreira, M., Fernandes, A.M., Aleixo, A., Antonelli, A., Olsson, U., Bates, J.M., ... and Ribas, C.C. (2018). Evidence for mtDNA capture in the jacamar Galbula leucogastra/chalcothorax species-complex and insights on the evolution of white-sand ecosystems in the Amazon basin. *Molecular phylogenetics and evolution*, 129, 149-157.

Filipski, A., Murillo, O., Freydenzon, A., Tamura, K., and Kumar, S. (2014). Prospects for building large time-trees using molecular data with incomplete gene coverage among species. *Molecular Biology and Evolution*, 31(9), 2542-2550.

Fischer, W. (Ed.). (1973). *FAO species identification sheets for fishery purposes: Mediterranean and Black Sea (fishing area 37) (Vol. 1)*. Rome, Italy: UN Food and Agriculture Organization (FAO).

Fischer, W., and Bianchi, G. (1984). FAO Species Identification Sheets for Fishery Purposes: Western Indian Ocean (Fishing Area 51). Volume II. Food and Agricultural Organization of the United Nations: Rome.

Floeter, S.R., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W.F., Wirtz, P., ... and Brito, A. (2008). Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, 35(1), 22-47.

Flower, B.P., and Kennett, J.P. (1993). Middle Miocene ocean-climate transition: High-resolution oxygen and carbon isotopic records from Deep Sea Drilling Project Site 588A, southwest Pacific. *Paleoceanography and Paleoclimatology*, 8(6), 811-843.

Flower, B.P., and Kennett, J.P. (1994). The middle Miocene climatic transition: East Antarctic ice sheet development, deep ocean circulation and global carbon cycling. *Palaeogeography, palaeoclimatology, palaeoecology*, 108(3-4), 537-555.

Foll, M., and Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, 180(2), 977-993.

Folland, C., Renwick, J., Salinger, M., Mullan, A. (2002). Relative influences of the Interdecadal Pacific Oscillation and ENSO on the South Pacific Convergence Zone. *Geophysical Research,* 29, 21-1-21-4.

Fraga, E., Schneider, H., Nirchio, M., Santa-Brigida, E., Rodrigues-Filho, L.F., and Sampaio, I. (2007). Molecular phylogenetic analyses of mullets (Mugilidae, Mugiliformes) based on two mitochondrial genes. *Journal of Applied Ichthyology*, 23(5), 598-604.

Franco, A., Pérez-Ruzafa, A., Drouineau, H., Franzoi, P., Koutrakis, E.T., Lepage, M., ... and Sapounidis, A. (2012). Assessment of fish assemblages in coastal lagoon habitats: Effect of sampling method. *Estuarine, Coastal and Shelf Science*, 112, 115-125.

Fratini, S., and Vannini, M. (2002). Genetic differentiation in the mud crab Scylla serrata (Decapoda : Portunidae) within the Indian Ocean. *Journal of Experimental Marine Biology and Ecology,* 272, 103-116.

Fratini, S., Ragionieri, L., and Cannicci, S. (2010). Stock structure and demographic history of the Indo-West Pacific mud crab Scylla serrata. *Estuarine Coastal and Shelf Science,* 86, 51-61.

Freyhof, J., Lieckfeldt, D., Pitra, C., and Ludwig, A. (2005). Molecules and morphology: evidence for introgression of mitochondrial DNA in Dalmatian cyprinids. *Molecular Phylogenetics and Evolution*, 37(2), 347-354.

Fricke, R., Eschmeyer, W.N., and Van der Laan R. (Eds.). 2019. CATALOG OF FISHES: GENERA, SPECIES, REFERENCES. In *Eschmeyer's Catalog of Fishes*. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (last accessed on 3 January 2019).

Froese, R., & Pauly, D. (Eds). (2001). Fish base online. Available at www.fishbase.org .

Fu, Y.X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147(2), 915-925.

Fuller, Z.L., Koury, S.A., Phadnis, N., Schaeffer, S.W. (2019) How chromosomal rearrangements shape adaptation and speciation: Case studies in Drosophila pseudoobscura and its sibling species Drosophila persimilis. *Molecular Ecology*. 28(6), 1283-1301

Funk, D.J., and Omland, K.E. (2003). Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, 34(1), 397-423.

Gagnaire, P.A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., ... & Bierne, N. (2015). Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary applications*, 8(8), 769-786.

Galtier, N., Enard, D., Radondy, Y., Bazin, E., and Belkhir, K. (2006). Mutation hot spots in mammalian mitochondrial DNA. *Genome Res*, 16, 215–222

Gammelsrød, T., Bartholomae, C.H., Boyer, D.C., Filipe, V.L.L., and O'toole, M.J. (1998). Intrusion of warm surface water along the Angolan Namibian Coast in February–March 1995: the 1995 Benguela niño. *African Journal of Marine Science*, 19.

García, F.J., and Bertsch, H. (2009). Diversity and distribution of the Gastropoda Opisthobranchia from the Atlantic Ocean: a global biogeographic approach. *Scientia Marina*, 73(1), 153-160.

García, M., Hernández-Molina, F.J., Llave, E., Stow, D.A.V., León, R., Fernández-Puga, M.C., ... and Somoza, L. (2009). Contourite erosive features caused by the Mediterranean Outflow Water in the Gulf of Cadiz: Quaternary tectonic and oceanographic implications. *Marine Geology*, 257(1-4), 24-40.

Garcia-Vazquez E., Machado-Schiaffino G., Campo D., and Juanes F. (2012). Species misidentification in mixed hake fisheries may lead to overexploitation and population bottlenecks. *Fisheries Research*, 114: 52–55.

Gardner, J.V., and Hays, J.D. (1976). The eastern equatorial Atlantic: sea-surface temperature and circulation responses to global climatic change during the past 200,000 years. In *Investigation of Late Quaternary Paleoceanography and Paleoclimatology* (Vol. 145, pp. 221-246).

Garnier, S., Alibert, P., Audiot, P., Prieur, B., and Rasplus, J.Y. (2004). Isolation by distance and sharp discontinuities in gene frequencies: implications for the phylogeography of an alpine insect species, Carabus solieri. *Molecular Ecology*, 13(7), 1883-1897.

Gascuel, D., Zeller, D., Taleb Sidi, M.O., and Pauly, D. (2007). Reconstructed catches in the Mauritanian EEZ. *Reconstruction of marine fisheries catches for key countries and regions (1950–2005). Vancouver (Canada): Fisheries Centre, University of British Columbia*, 105-119.

George, R.W. (2006). Tethys sea fragmentation and speciation of Panulirus spiny lobsters. *Crustaceana*, 78 (11), 1281-1309.

Geyer, L.B., and Palumbi, S.R. (2003). Reproductive character displacement and the genetics of gamete recognition in tropical sea urchins. *Evolution*, 57(5), 1049-1060.

Ghasemzadeh, J., (1998). Phylogeny and Systematics of Indo-Pacific mullets (Teleostei: Mugilidae) with Special Reference to the Mullets of Australia. Ph.D. Dissertation. Macquarie University, Sydney, 397pp.

Gienapp, P., Teplitsky, C., Alho, J.S., Mills, J.A., and Merilä, J. (2008). Climate change and evolution: disentangling environmental and genetic responses. *Molecular ecology*, 17(1), 167-178.

Gillooly, J.F., Allen, A.P., West, G.B., and Brown, J.H. (2005). The rate of DNA evolution: effects of body size and temperature on the molecular clock. *Proceedings of the National Academy of Sciences*, 102(1), 140-145.

Gilpin, R. (2007). *Enhancing maritime security in the Gulf of Guinea*. Naval Postgraduate School, Monterey CA Centre for Contemporary Conflict.

Gittenberger, A., and Schrieken, N. (2004). Octopuses, squids and cuttlefish (Cephalopoda) in the Netherlands. *Vita Malacologica*, 2: 33–38.

Glynn, F., Houghton, J.D.R., and Provan, J. (2015). Population genetic analyses reveal distinct geographical blooms of the jellyfish Rhizostoma octopus (Scyphozoa). *Biological Journal of the Linnean Society,* 116, 582-592.

Glynn, F., Houghton, J.D.R., Bastian, T., Doyle, T.K., Fuentes, V., Lilley, M.K.S., and Provan, J. (2016). High-resolution genetic analysis reveals extensive gene flow within the jellyfish Pelagia noctiluca (Scyphozoa) in the North Atlantic and Mediterranean Sea. *Biological Journal of the Linnean Society,* 117, 252-263.

Goff, R.L., and Daguzan, J. (1991). Growth and life cycles of the cuttlefish Sepia officinalis L.(Mollusca: Cephalopoda) in South Brittany (France). *Bulletin of Marine Science*, 49(1-2), 341-348.

Golikov, A.V., Sabirov, R.M., Lubin, P.A., Jørgensen, L.L., and Beck, I.M. (2014). The northernmost record of Sepietta oweniana (Cephalopoda: Sepiolidae) and comments on boreo-subtropical cephalopod species occurrence in the Arctic. *Marine Biodiversity Records*, *7*.

Gonzalez, E.G., and Zardoya, R. (2007). Relative role of life-history traits and historical factors in shaping genetic population structure of sardines (Sardina pilchardus). *BMC evolutionary biology*, 7(1), 197.

Gonzalez-Wevar, C.A., David, B., and Poulin, E. (2011). Phylogeography and demographic inference in *Nacella*

(Patinigera) concinna (Strebel, 1908) in the western Antarctic Peninsula. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(1-2), 220-229.

Good, J.M., Hird, S., Reid, N., Demboski, J.R., Steppan, S.J., Martin-Nims, T.R., and Sullivan, J. (2008). Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology*, 17 (5), 1313-1327.

Good, J.M., Vanderpool, D., Keeble, S., and Bi, K. (2015). Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. *Evolution*, 69(8), 1961-1972.

Gopal, K., Tolley, K.A., Groeneveld, J.C., and Matthee, C.A. (2006). Mitochondrial DNA variation in spiny lobster Palinurus delagoae suggests genetically structured populations in the southwestern Indian Ocean. *Marine Ecology Progress Series*, 319, 191-198.

Gornung, E., Colangelo, P., and Annesi, F. (2007). 5S ribosomal RNA genes in six species of Mediterranean grey mullets: genomic organization and phylogenetic inference. *Genome*, 50(9), 787-795.

Goudet, J. (1995). FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of heredity*, 86 (6), 485-486.

Gourène, A.B., Pouyaud, L., and Agnèse, J.F. (1993). Importance de certaines caractéristiques biologiques dans la structuration génétique des especes de poissons: le cas de Ethmalosa fimbriata et Sarotherodon melanotheron. *Journal Ivoirien d'Océanologie et de Limnologie*, 2(2), 55-69.

Gouws G. (2012). Genetic stock assessment of the blue emperor (Lethrinus nebulosus) in the South West Indian Ocean. *South African Institute for Aquatic Biodiversity National Research Foundation*. 1-32.

Grabowski, M., Grater, W.D., and Stuck, K.C. (2004). A novel polymorphic mtDNA marker for population studies of the pink shrimp, Farfantepenaeus duorarum (Crustacea, Penaeidae). *Oceanologia*, 46(1).

Gradstein, F.M., Ogg, J.G., Schmitz, M.D., and Ogg, G.M. (2012). The Geological Time Scale 2012 (p. 1144). Amsterdam: Elsevier

Grandcourt, E.M. (2002). Demographic characteristics of a selection of exploited reef fish from the Seychelles: preliminary study. *Marine and Freshwater Research,* 53, 123-130.

Grandcourt, E.M., Al Abdessalaam, T.Z., Al Shamsi, A.T., and Francis, K. (2006). Biology and assessment of the painted sweetlips (*Diagramma pictum* (Thunberg, 1792)) and the spangled emperor (*Lethrinus nebulosus* (Forsskål, 1775)) in the southern Arabian Gulf. *Fisheries Bulletin*, 104**:** 75–88.

Grant, P.R., and Grant, B.R. (1997). Genetics and the origin of bird species. *Proceedings of the National Academy of Sciences*, 94(15), 7768-7775..

Grant, W.S. (2015). Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography. *Journal of Heredity*, 106(4), 333-346.

Grant, W.S., and Bowen, B.W. (2006). Living in a tilted world: climate change and geography limit speciation in Old World anchovies (Engraulis; Engraulidae). *Biological Journal of the Linnean Society*, 88(4), 673-689.

Grant, W.S., and Leslie, R.W. (2001). Inter-ocean dispersal is an important mechanism in the zoogeography of hakes (Pisces: Merluccius spp.). *Journal of Biogeography*, 28(6), 699-721.

Grant, W.S., Leslie, R.W., and Bowen, B.W. (2005). Molecular genetic assessment of bipolarity in the anchovy genus Engraulis. *Journal of Fish Biology*, 67(5), 1242-1265.

Griffiths, A.M., Sims, D.W., Johnson, A., Lynghammar, A., McHugh, M., Bakken, T., and Genner, M.J. (2011). Levels of connectivity between longnose skate (Dipturus oxyrinchus) in the Mediterranean Sea and the north-eastern Atlantic Ocean. *Conservation Genetics*, 12(2), 577-582.

Griffiths, M.H., and Heemstra, P.C. (1995). *A contribution to the taxonomy of the marine fish genus Argyrosomus (Perciformes: Sciaenidae), with descriptions of two new species from southern Africa*. Icthyological Bulletin, 65, 1-40

Guerra, A., Pérez-Losada, M., Rocha, F., and Sanjuan, A. (2001). Species differentiation of *Sepia officinalis* and *Sepia hierredda* (Cephalopoda: Sepiidae) based on morphological and allozyme analyses. *Journal of the Marine Biological Association of the UK*, 81: 271–281.

Guillemot, N., Léopold, M., Cuif, M., and Chabanet, P. (2009). Characterization and management of informal fisheries confronted with socio-economic changes in New Caledonia (South Pacific). *Fisheries research*, 98(1-3), 51-61.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML

3.0. *Systematic biology*, 59(3), 307-321.

Gwilliam, M. (2017). *Assessing the role of the Benguela Current as a major biogeographic barrier to marine coastal species* (Doctoral dissertation, Aberystwyth University)

Gwilliam, M.P., Winkler, A.C., Potts, W.M., Santos, C.V., Sauer, W.H.H., Shaw, P.W., and McKeown, N. J. (2018). Integrated genetic and morphological data support eco-evolutionary divergence of Angolan and South African populations of Diplodus hottentotus. *Journal of fish biology*, 92(4), 1163-1176.

Hagen, E., Zulicke, C., and Feistel, R. (1996). Near-surface structures in the Cape Ghir filament off Morocco. *Oceanologica Acta*, 19(6), 577-598.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series,* 41, 95-98.

Han, Y.S., Hung, C.L., Liao, Y.F., & Tzeng, W.N. (2010). Population genetic structure of the Japanese eel Anguilla japonica: panmixia at spatial and temporal scales. *Marine Ecology Progress Series*, 401, 221-232.

Hansen, M.M., Meier, K., and Mensberg, K.L.D. (2010). Identifying footprints of selection in stocked brown trout populations: a spatio-temporal approach. *Molecular Ecology*, 19(9), 1787-1800.

Hardman-Mountford, N.J., and McGlade, J.M. (2003). Seasonal and interannual variability of oceanographic processes in the Gulf of Guinea: an investigation using AVHRR sea surface temperature data. *International Journal of Remote Sensing*, 24(16), 3247-3268.

Harrison, I.J. (2007). Mugilidae. In: Stiassny MLJ, Teugels GG, Hopkins CD, editors. Poissons d'eaux douces et saumâtres de basse Guinée, ouest de l'Afrique centrale: Faune et flore tropicales. Paris: Institut de Recherche pour le Développement, pp. 450–471.

Harrison, I.J., and Howes, G.J. (1991). The pharyngobranchial organ of mugilid fishes; its structure, variability, ontogeny, possible function and taxonomic utility. *Bulletin of the British Museum, Natural History. Zoology*, 57(2), 111-132.

Harrison, I.J., and Senou, H. (1999) Mugilidae. In: Carpenter KE, Niem VH (eds) FAO species identification guide for fishery purposes. The living marine resources of the Western Central Pacific, vol 4: bony fishes part 2 (Mugilidae to Carangidae). FAO, Rome, pp 2069–2108

Harrison, T.D. (2002). Preliminary assessment of the biogeography of fishes in South African estuaries. *Marine and Freshwater Research*, 53(2), 479-490.

Harrison, T.D. (2003). *Biogeography and community structure of fishes in South African estuaries* (Doctoral dissertation, Rhodes University).

Harrison, T.D., and Whitfield, A.K. (2006). Temperature and salinity as primary determinants influencing the biogeography of fishes in South African estuaries. *Estuarine, Coastal and Shelf Science*, 66(1-2), 335-345.

Hassan, M., Lemaire, C., Fauvelot, C., & Bonhomme, F. (2002). Seventeen new exon-primed intron-crossing polymerase chain reaction amplifiable introns in fish. *Molecular Ecology Notes*, 2(3), 334-340.

Hatanaka, H. (ed.). (1979). Geographical distribution of two subspecies of *Sepia officinalis* LINNE off the northwestern coast of Africa. In: Report of the ad hoc Working Group on the Assessment of Cephalopod Stocks, Appendix 4. Fishery Committee for the Eastern Central Atlantic CECAF/ECAF SERIES 78/11.

Hauser, L., and Carvalho, G.R. (2008). Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries*, 9(4), 333-362.

Hay, W.W., and Brock, J.C. (1992). Temporal variation in intensity of upwelling off southwest Africa. *Geological Society, London, Special Publications*, 64(1), 463-497.

Haynes, R., Barton, E.D., and Pilling, I. (1993). Development, persistence, and variability of upwelling filaments off the Atlantic coast of the Iberian Peninsula. *Journal of Geophysical Research: Oceans*, 98 (C12), 22681-22692.

Hazin, F.H.V. (1993). *Fisheries-oceanographical study on tunas, billfishes and sharks in the Southwestern Equatorial Atlantic Ocean*. (Doctoral dissertation, Tokyo University of Fisheries).

Healey, A. J., Gouws, G., Fennessy, S. T., Kuguru, B., Sauer, W. H., Shaw, P. W., ... and Handling editor: W. Stewart Grant. (2018b). Genetic analysis reveals harvested Lethrinus nebulosus in the Southwest Indian Ocean comprise two cryptic species. *ICES Journal of Marine Science*, *75*(4), 1465-1472.

Healey, A.J., McKeown, N.J., Taylor, A.L., Provan, J., Sauer, W., Gouws, G., and Shaw, P.W. (2018a). Cryptic species and parallel genetic structuring in Lethrinid fish: Implications for conservation and management in the southwest Indian Ocean. *Ecology and evolution*, *8*(4), 2182-2195.

Healey, A.J.E., McKeown, N.J., Potts, W.M., de Beer, C.L., Sauer, W., and Shaw, P.W. (2017). Phylogeny of the Sepia officinalis species complex in the eastern Atlantic extends the known distribution of Sepia vermiculata across the Benguela upwelling region. *African journal of marine science*, 39(3), 307-313.

Heath, T.A., Huelsenbeck, J.P., and Stadler, T. (2014). The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences*, 111(29), E2957-E2966.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., and Francis, C.M. (2004). Identification of birds through DNA barcodes. *PloS Biology*, 2: 1657–1663.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., and Francis, C.M. (2004). Identification of birds through DNA barcodes. *PloS Biology*, 2: 1657–1663.

Hecht, T. (1990). On the life history of Cape horse mackerel *Trachurus trachurus capensis* off the south-east coast of South Africa. *South African Journal of Marine Science*, 9(1), 317-326.

Hedrick, P.W. (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular ecology*, *22*(18), 4606-4618.

Heemstra, P.C., and Heemstra, E. (2004). *Coastal fishes of southern Africa*. NISC (PTY) LTD.

Heileman, S., Fennessy, S.T., and van der Elst, R.P. (2015). Demersal fisheries, a retrospective analysis of their status in the Southwest Indian Ocean. *In* Offshore fisheries of the Southwest Indian Ocean: their status and the impact on vulnerable species. pp. 213–285. Ed. by R. P. Van der Elst, and B. I. Everett. Oceanographic Research Institute, Special Publication, 10. 448pp.

Heinrich, S., Zonneveld, K.A., Bickert, T., and Willems, H. (2011). The Benguela upwelling related to the Miocene cooling events and the development of the Antarctic Circumpolar Current: Evidence from calcareous dinoflagellate cysts. *Paleoceanography*, 26(3).

Helyar, S.J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M.I., Ogden, R., Limborg, M.T., ... & Nielsen, E.E. (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular ecology resources*, 11, 123-136.

Helyar, S.J., Lloyd, H.A.D., de Bruyn, M., Leake, J., Bennett, N., and Carvalho, G.R. (2014). Fish product mislabelling: failings of traceability in the production chain and implications for illegal, unreported and unregulated (IUU) fishing. *PLoS One*, 9(6), e98691.

Hemmer-Hansen, J., Hüssy, K., Baktoft, H., Huwer, B., Bekkevold, D., Haslob, H., ... & Nielsen, E.E. (2019). Genetic analyses reveal complex dynamics within a marine fish management area. *Evolutionary applications*, 12(4), 830-844.

Hemmer-Hansen, J., Nielsen, E.E., Frydenberg, J., and Loeschcke, V. (2007). Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus L.*). *Heredity*, 99 (6): 592–600.

Henin, C., P. Hisard, and B. Piton, (1986). Les eaux de l'Atlantique equatorial pendant l'op6ration FOCAL (juillet 1982-ao6t 1984), Tray. Doc. ORSTOM, 196 pp.

Henke, D., Smyth, P., Haffke, C., and Magnusdottir, G. (2012). Automated analysis of the temporal behavior of the double Intertropical Convergence Zone over the east Pacific. *Remote sensing of environment*, 123, 418-433.

Henkel, S.K., and Hofmann, G.E. (2008). Differing patterns of hsp70 gene expression in invasive and native kelp species: evidence for acclimation-induced variation. *Journal of Applied Phycology*, 20(5), 915-924.

Henriques, R. (2012). *Influence of the Benguela Current in genetic sub-structuring of commercially exploited fish species* (Doctoral dissertation, Royal Holloway, University of London).

Henriques, R., Potts, W.M., Santos, C.V., Sauer, W.H., and Shaw, P.W. (2014). Population connectivity and phylogeography of a coastal fish, Atractoscion aequidens (Sciaenidae), across the Benguela Current Region: evidence of an ancient vicariant event. *PloS one*, 9(2), e87907.

Henriques, R., Potts, W.M., Santos, C.V., Sauer, W.H.H., and Shaw, P.W. (2018). Population connectivity of an overexploited coastal fish, Argyrosomus coronus (Sciaenidae), in an ocean-warming hotspot. *African Journal of Marine Science*, 40(1), 13-24.

Henriques, R., Potts, W.M., Sauer, W.H., Santos, C.V., Kruger, J., Thomas, J.A., and Shaw, P.W. (2016). Molecular genetic, life-history and morphological variation in a coastal warm-temperate sciaenid fish: evidence for an upwelling-driven speciation event. *Journal of biogeography*, 43(9), 1820-1831.

Henriques, R., Potts, W.M., Sauer, W.H.H., and Shaw, P.W. (2012). Evidence of deep genetic divergence between populations of an important recreational fishery species, Lichia amia L. 1758, around southern

Africa.*African Journal of Marine Science*, 34(4), 585-591.

Henriques, R., von der Heyden, S., Lipinski, M.R., du Toit, N., Kainge, P., Bloomer, P., and Matthee, C.A. (2016). Spatio-temporal genetic structure and the effects of long-term fishing in two partially sympatric offshore demersal fishes. *Molecular ecology*, 25(23), 5843-5861.

Heras, S., Castro, M.G., and Roldán, M.I. (2006). Mugil curema in Argentinean waters: combined morphological and molecular approach. *Aquaculture*, 261(2), 473-478.

Heras, S., Roldán, M.I., and Castro, M.G. (2009). Molecular phylogeny of Mugilidae fishes revised. *Reviews in Fish Biology and Fisheries*, 19(2), 217-231.

Hernández-Guerra, A., and Nykjaer, L. (1997). Sea surface temperature variability off north-west Africa: 1981 -1989. *International Journal of Remote Sensing*, 18(12), 2539-2558.

Hess, J.E., Campbell, N.R., Close, D.A., Docker, M.F., and Narum, S.R. (2013). Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. *Molecular Ecology*, 22(11), 2898-2916.

Hey, J. (2009). Isolation with migration models for more than two populations. *Molecular biology and evolution*, 27(4), 905-920.

Hey, J., and Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, 167(2), 747-760.

Hickerson, M.J., and Meyer, C.P. (2008). Testing comparative phylogeographic models of marine vicariance and dispersal using a hierarchical Bayesian approach. *BMC evolutionary biology*, 8(1), 322.

Hipsley, C.A., and Müller, J. (2014). Beyond fossil calibrations: realities of molecular clock practices in evolutionary biology. *Frontiers in genetics*, 5, 138.

Ho, S.Y., and Phillips, M.J. (2009). Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, 58(3), 367-380.

Ho, S.Y., Lanfear, R., Bromham, L., Phillips, M.J., Soubrier, J., Rodrigo, A.G., and Cooper, A. (2011). Time-dependent rates of molecular evolution. *Molecular ecology*, 20(15), 3087-3101.

Ho, S.Y., Phillips, M.J., Cooper, A., and Drummond, A.J. (2005). Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular biology and evolution*, 22(7), 1561-1568.

Ho, S.Y., Shapiro, B., Phillips, M.J., Cooper, A., and Drummond, A.J. (2007). Evidence for time dependency of molecular rate estimates. *Systematic biology*, 56(3), 515-522.

Hoarau, G., Rijnsdorp, A.D., Van der Veer, H.W., Stam, W.T., & Olsen, J.L. (2002). Population structure of plaice (Pleuronectes platessa L.) in northern Europe: microsatellites revealed large-scale spatial and temporal homogeneity. *Molecular Ecology*, 11(7), 1165-1176.

Hoareau, T.B. (2015). Late glacial demographic expansion motivates a clock overhaul for population genetics. *Systematic biology*, 65(3), 449-464.

Hoareau, T.B., Boissin, E., and Berrebi, P. (2012). Evolutionary history of a widespread Indo-Pacific goby: the role of Pleistocene sea-level changes on demographic contraction/expansion dynamics. *Molecular phylogenetics and evolution*, 62(1), 566-572.

Hoareau, T.B., Boissin, E., Paulay, G., and Bruggemann, J.H. (2013). The Southwestern Indian Ocean as a potential marine evolutionary hotspot: perspectives from comparative phylogeography of reef brittlestars. *Journal of Biogeography*, 40(11), 2167-2179.

Hodell, D.A., and Venz, K. (1992). Toward a high-resolution stable isotopic record of the Southern Ocean during the Pliocene-Pleistocene (4.8 to 0.8 Ma). *The Antarctic Paleoenvironment: A Perspective on Global Change: Part One*, 56, 265-310.

Hodell, D.A., Benson, R.H., Kent, D.V., Boersma, A., and Bied, R.E. (1994). Magnetostratigraphic, biostratigraphic, and stable isotope stratigraphy of an Upper Miocene drill core from the Salé Briqueterie (northwestern Morocco): A high-resolution chronology for the Messinian stage. *Paleoceanography and Paleoclimatology*, 9(6), 835-855.

Hoegh-Guldberg, O., & Bruno, J.F. (2010). The impact of climate change on the world's marine ecosystems. *Science*, 328(5985), 1523-1528.

Hoffmann, A. A., and Sgro, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*(7335), 479.

Holbourn, A., Kuhnt, W., Clemens, S., Prell, W., and Andersen, N. (2013). Middle to late Miocene stepwise

climate cooling: Evidence from a high-resolution deep water isotope curve spanning 8 million years. *Paleoceanography*, 28(4), 688-699.

Hooper, D.M. (2016). Range overlap drives chromosome inversion fixation in passerine birds. *bioRxiv*, 053371.

Hooper, D.M., Griffith, S.C., and Price, T.D. (2019). Sex chromosome inversions enforce reproductive isolation across an avian hybrid zone. *Molecular ecology*, 28(6), 1246-1262.

Horne, J.B., van Herwerden, L., Choat, J.H., and Robertson, D.R. (2008). High population connectivity across the Indo-Pacific: congruent lack of phylogeographic structure in three reef fish congeners. *Molecular phylogenetics and evolution*, 49(2), 629-638.

Houghton, R.W., and Colin, C. (1986). Thermal structure along 4 W in the Gulf of Guinea during 1983–1984. *Journal of Geophysical Research: Oceans*, 91(C10), 11727-11739.

Hovan, S.A. (1995). 28. Late Cenozoic atmospheric circulation intensity and climatic history recorded by Eolian deposition in the Eastern Equatorial Pacific Ocean, Leg 138. *Proceedings of the Ocean Drilling Program, Scientific Results. Proceedings of the Ocean Drilling Program, Scientific Results*, 615-625.

Hrbek, T., and Meyer, A. (2003). Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). *Journal of Evolutionary Biology*, 16(1), 17-36.

Hsü, K.J. (1973). The desiccated deep-basin model for the Messinian events. *Messinian events in the Mediterranean*, 60-67.

Hubbs, C.L. (1952). Antitropical distribution of fishes and other organisms. Symposium on problems of bipolarity and of pantemperate faunas. *Proceedings of the Seventh Pacific Science. Congress .(Pac. Sci. Assoc.)*, 3, 324-329.

Hubisz, M.J., Falush, D., Stephens, M., and Pritchard, J.K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular ecology resources*, 9(5), 1322-1332.

Hutchings, L., Van der Lingen, C.D., Shannon, L.J., Crawford, R.J.M., Verheye, H.M.S., Bartholomae, C.H., ... and Fidel, Q. (2009). The Benguela Current: An ecosystem of four components. *Progress in Oceanography*, 83(1-4), 15-32.

Hutchinson, J.A., and Reynolds, J.D. (2004). Marine fish population collapses: consequences for recovery and extinction risk. *Bioscience*, 54: 297–309

Huyse, T., Van Houdt, J., and Volckaert, F.A. (2004). Paleoclimatic history and vicariant speciation in the "sand goby" group (Gobiidae, Teleostei). *Molecular Phylogenetics and Evolution, 32(*1), 324-336.

Iacchei, M., Ben-Horin, T., Selkoe, K.A., Bird, C.E., García-Rodríguez, F.J., and Toonen, R.J. (2013). Combined analyses of kinship and FST suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. *Molecular Ecology*, 22(13), 3476-3494.

ICES (2005) Report of the working group on the assessment of mackerel, horse mackerel, sardine and anchovy (WGMHSA). ICES CM 2005/ACFM: 08.

ICES, (2003). Report of the Working Group on the Assessment of Mackerel, Horse Mackerel, Sardine and Anchovy. ICES C.M. 2003/ACFM: 07.

Idyll, C.P., and Sutton, J.W. (1952). Results of the first year's tagging of mullet, Mugil cephalus L., on the west coast of Florida. *Transactions of the American Fisheries Society*, 81(1), 69-77.

Iles, T.D., and Sinclair, M. (1982). Atlantic herring: stock discreteness and abundance. *Science*, 215(4533): 627–633.

Imsiridou, A., Minos, G., Katsares, V., Karaiskou, N., and Tsiora, A. (2007). Genetic identification and phylogenetic inferences in different Mugilidae species using 5S rDNA markers. *Aquaculture Research*, 38(13), 1370-1379.

Ingham, M.C. (1970). Coastal upwelling in the northwestern Gulf of Guinea. *Bulletin of Marine Science*, 20(1), 1-34.

Innal, D., and Erk'akan, F. (2006). Effects of exotic and translocated fish species in the inland waters of Turkey. *Reviews in Fish Biology and Fisheries*, 16(1), 39-50.

Jahn, B., Donner, B., Müller, P.J., Röhl, U., Schneider, R.R., and Wefer, G. (2003). Pleistocene variations in dust input and marine productivity in the northern Benguela Current: evidence of evolution of global glacial–interglacial cycles. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 193(3), 515-533.

Jamandre, B.W., Durand, J.D., and Tzeng, W.N. (2009). Phylogeography of the flathead mullet Mugil cephalus

in the north-west Pacific as inferred from the mtDNA control region. *Journal of fish biology*, 75(2), 393 -407.

James, N.C., Whitfield, A.K., and Harrison, T.D. (2016). Grey mullet (Mugilidae) as possible indicators of global warming in South African estuaries and coastal waters. *Marine environmental research*, 122, 188-195.

Janko, K., Lecointre, G., DeVries, A., Couloux, A., Cruaud, C., and Marshall, C. (2007). Did glacial advances during the Pleistocene influence differently the demographic histories of benthic and pelagic Antarctic shelf fishes?–Inferences from intraspecific mitochondrial and nuclear DNA sequence diversity. *BMC Evolutionary Biology*, 7(1), 220.

Jaziri, H., and Benazzou, T. (2002). Multilocus Allozymic Differentiation of Mold Populations Mytilus galloprovincialis Lmk. Moroccan coasts. *Accounts Biology,* 325 (12), 1175-1183

Jeffreys, H. (1961). Theory of Probability. *Oxford University Press, Oxford, UK.*

Jereb, P., Allcock, A.L., Lefkaditou, E., Piatkowski, U., Hastie, L.C., and Pierce, G.J. (eds). (2015). Cephalopod biology and fisheries in Europe: II. Species Accounts. *ICES Cooperative Research Report* No. 325: 360.

Jermiin, L.S., Graur, D., Lowe, R.M., & Crozier, R.H. (1994). Analysis of directional mutation pressure and nucleotide content in mitochondrial cytochrome b genes. *Journal of Molecular Evolution*, 39(2), 160-173.

Jerry, D.R., & Baverstock, P.R. (1998). Consequences of a catadromous life-strategy for levels of mitochondrial DNA differentiation among populations of the Australian bass, Macquaria novemaculeata. *Molecular Ecology*, 7(8), 1003-1013.

Jiang, S., Wise, S.W., Wang, Y., and Teagle, D.A.H. (2007). Cause of the middle/late Miocene carbonate crash: dissolution or low productivity. In *Proceedings of the Ocean Drilling Program, Scientific Results* (Vol. 206, pp. 1-24). College Station, TX: Ocean Drilling Program.

Jollit, I., Léopold, M., Guillemot, N., David, G., Chabanet, P., Lebigre, J.M., and Ferraris, J. (2010). Geographical aspects of informal reef fishery systems in New Caledonia. *Marine Pollution Bulletin*, 61(7-12), 585 -597.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24 (11), 1403-1405.

Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal component: a new method for the analysis of genetically structure populations. *BMC Genetics*, 11, 94.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., ... and Birney, E. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, 484(7392), 55.

Kalinowski, S.T., Wagner, A.P., and Taper, M.L. (2006). ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, 6(2), 576-579.

Kapun, M., and Flatt, T. (2019) The adaptive significance of chromosomal inversion polymorphisms in Drosophila melanogaster. *Molecular Ecology*. 28(6), 1263-1282

Karaiskou, N., Apostolidis, A.P., Triantafyllidis, A., Kouvatsi, A., and Triantaphyllidis, C. (2003). Genetic identification and phylogeny of three species of the genus *Trachurus* based on mitochondrial DNA analysis. *Marine Biotechn*ology, 5(5), 493-504.

Karaiskou, N., Triantafyllidis, A., and Triantaphyllidis, C. (2004). Shallow genetic structure of three species of the genus Trachurus in European waters. *Marine Ecology Progress Series*, 281, 193-205.

Kasapidis, P., and Magoulas, A. (2008). Development and application of microsatellite markers to address the population structure of the horse mackerel *Trachurus trachurus*. *Fisheries Research*, 89(2), 132-135.

Kawashima, S., and Kanehisa, M. (2000). AAindex: amino acid index database. *Nucleic acids research*, 28(1), 374-374.

Kay, E.A., and Palumbi, S.R. (1987). Endemism and evolution in Hawaiian marine invertebrates. *Trends in Ecology and Evolution*, 2(7), 183-186.

Ke, H.M., Lin, W.W., and Kao, H.W. (2009). Genetic diversity and differentiation of gray mullet (Mugil cephalus) in the coastal waters of Taiwan. *Zoological science*, 26(6), 421-428.

Kelley, C.D., Tamaru, C.S., Lee, C.S., Moriwake, A., and Miyamoto, G. (1991). Effects of photoperiod and temperature on the annual ovarian cycle of the striped mullet, Mugil cephalus.
A.P. Scott, J.P. Sumpter, D.E. Kime, M.S. Rolfe (Eds.), Proceedings of the Fourth International Symposium on the Reproductive Physiology of Fish, FishSymp, Sheffield (1991), pp. 142-144

Keskİn, E., and Atar, H.H. (2013). DNA barcoding commercially important fish species of T urkey. *Molecular ecology resources*, 13(5), 788-797.

Khromov, D.N. (1998). Distribution patterns of Sepiidae. *Smithsonian Contribution to Zoology*, 586: 191–206.

Khromov, D.N., Lu, C.C., Guerra, A., Dong, Z., and Boletzky, S.V. (1998). A synopsis of Sepiidae outside Australian waters (Cephalopoda: Sepioidea). *Smithsonian Contribution to Zoology*, 586: 77–139.

Kimani, E.N., Okemwa, G.M., and Kazungu, J.M. (2009). Fisheries in the Southwest Indian Ocean: trends and governance challenges. In The Indian Ocean: Resource and Governance Challenges, Laipson E, Pandya A (eds). The Henry L. Stimson Centre: Washington DC.

Kimura, M., & Weiss, G.H. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, 49(4), 561.

King, N.G., McKeown, N.J., Smale, D.A., and Moore, P.J. (2018). The importance of phenotypic plasticity and local adaptation in driving intraspecific variability in thermal niches of marine macrophytes. *Ecography*, 41(9), 1469-1484.

King, N.G., McKeown, N.J., Smale, D.A., Wilcockson, D.C., Hoelters, L., Groves, E.A., ... and Moore, P.J. (2019). Evidence for different thermal ecotypes in range centre and trailing edge kelp populations. *Journal of Experimental Marine Biology and Ecology*, 514, 10-17.

Kirchner, C., Bauleth-D'Almeida, G., and Wilhelm, M.R. (2010). Assessment and management of Cape horse mackerel *Trachurus capensis* off Namibia based on a fleet-disaggregated age-structured production model. *African Journal of Marine Science*, 32(3), 525-541.

Kirkpatrick, M., and Ravigné, V. (2002). Speciation by natural and sexual selection: models and experiments. *the american naturalist*, 159(S3), S22-S35.

Kitazaki, K., and Kubo, T. (2010). Cost of having the largest mitochondrial genome: evolutionary mechanism of plant mitochondrial genome. *Journal of Botany*.

Klein, B., and Siedler, G. (1989). On the origin of the Azores Current. *Journal of Geophysical Research: Oceans*, 94(C5), 6159-6168.

Knight, A.J., Hughes, R.N., and Ward, R.D. (1987). A striking example of the founder effect in the mollusc Littorina saxatilis. *Biological Journal of the Linnean Society*, 32(4), 417-426.

Knijn, R.J., Boon, T.W., Heessen, H.J., and Hislop, J.R. (1993). Atlas of North Sea fishes. *ICES cooperative research report*, 194, 268.

Knoll, M., Hernández-Guerra, A., Lenz, B., Laatzen, F.L., Machın, F., Müller, T.J., and Siedler, G. (2002). The eastern boundary current system between the Canary Islands and the African coast. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(17), 3427-3440.

Knowlton, N., and Weigt, L.A. (1998). New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1412), 2257-2263.

Knutsen, H., Olsen, E.M., Jorde, P.E., Espeland, S.H., André, C., and Stenseth, N.C. (2011). Are low but statistically significant levels of genetic differentiation in marine fishes 'biologically meaningful'? A case study of coastal Atlantic cod. *Molecular Ecology*, 20(4): 768–783.

Koranteng, K.A. (2002). Fish Species Assemblages on the Continental Shelf and Upper slope off Ghana. In J.M. McGlade, P. Cury, K.A. Koranteng and N.J. Hardman-Mountford (Eds.), *The Gulf of Guinea Large Marine Ecosystem*: *Environmental Forcing and Sustainable Development of Marine Resources* (pp. 173–87). Amsterdam, Netherlands: Elsevier.

Koranteng, K.A., and McGlade, J.M. (2001). Climatic trends in continental shelf waters off Ghana and in the Gulf of Guinea, 1963–1992. *Oceanologica Acta*, 24(2), 187-198.

Koranteng, K.A., Ofori-Danson, P.K., and Entsua-Mensah, M. (2000). Fish and fisheries of the Muni lagoon in Ghana, West Africa. *Biodiversity and Conservation*, 9(4), 487-499.

Kotoulas, G., Magoulas, A., Tsimenides, N., and Zouros, E. (1995). Marked mitochondrial DNA differences between Mediterranean and Atlantic populations of the swordfish, Xiphias gladius. *Molecular Ecology*, 4(4), 473-482.

Kozak, G.M., Reisland, M., and Boughmann, J.W. (2009). Sex differences in mate recognition and conspecific preference in species with mutual mate choice. *Evolution: International Journal of Organic Evolution*, 63(2), 353-365.

Krammer, R., Baumann, K.H., and Henrich, R. (2006). Middle to late Miocene fluctuations in the incipient Benguela Upwelling System revealed by calcareous nannofossil assemblages (ODP Site 1085A). *Palaeogeography, Palaeoclimatology, Palaeoecology*, 230(3-4), 319-334.

Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., and Wilson, D.S. (1999). Chronology, causes and progression of the Messinian salinity crisis. *Nature*, 400(6745), 652.

Kulmiye, A.J., Ntiba, M.J., and Kisia, S.M. (2002). Some Aspects of the Reproductive Biology of the Thumbprint Emperor, Lethrinus harak (Forsskal, 1775), in Kenyan Coastal Waters. *Western Indian Ocean Journal of Marine Science*, 1, 135-144.

Kumazawa, Y., Ota, H., Nishida, M., and Ozawa, T. (1998). The complete nucleotide sequence of a snake (Dinodon semicarinatus) mitochondrial genome with two identical control regions. *Genetics*, 150(1), 313-329.

Kuo, C.H., and Avise, J.C. (2005). Phylogeographic breaks in low-dispersal species: the emergence of concordance across gene trees. *Genetica*, 124(2-3), 179-186.

Kuo, C.M. (1995). Manipulation of ovarian development and spawning in grey mullet, Mugil cephalus L. *Israeli Journal of Aquaculture-Bamidgeh*, 47(2), 43-58.

Kuparinen, A., Keith, D.M., and Hutchings, J.A. (2014). Allee effect and the uncertainty of population recovery. *Conservation Biology*, 28(3): 790-798.

Laconcha, U., Iriondo, M., Arrizabalaga, H., Manzano, C., Markaide, P., Montes, I., ... & Santiago, J. (2015). New nuclear SNP markers unravel the genetic structure and effective population size of albacore tuna (Thunnus alalunga). *PLoS One*, 10(6), e0128247.

Ladner, J.T., and Palumbi, S.R. (2012). Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes. *Molecular Ecology*, 21(9), 2224-2238.

Laikre, L., Schwartz, M.K., Waples, R.S., Ryman, N., and GeM Working Group. (2010). Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals. *Trends in ecology and evolution*, 25(9), 520-529.

Lamaze, F.C., Sauvage, C., Marie, A., Garant, D., and Bernatchez, L. (2012). Dynamics of introgressive hybridization assessed by SNP population genomics of coding genes in stocked brook charr (Salvelinus fontinalis). *Molecular Ecology*, 21(12), 2877-2895.

Lambeck, K., Esat, T.M., and Potter, E.K. (2002). Links between climate and sea levels for the past three million years. *Nature*, 419(6903), 199.

Larson, H.K., Williams, R.S., and Hammer, M.P. (2013). An annotated checklist of the fishes of the Northern Territory, Australia. *Zootaxa*, 3696(1), 1-293.

Larson, R.J., and Julian, R.M. (1999). Spatial and temporal genetic patchiness in marine populations and their implications for fisheries management. *California Cooperative Oceanic Fisheries Investigations Reports,* 40, 94-99.

Larsson, L.C., Charlier, J., Laikre, L., and Ryman, N. (2009). Statistical power for detecting genetic divergence—organelle versus nuclear markers. *Conservation Genetics*, 10(5), 1255.

Lass, H.U., Schmidt, M., Mohrholz, V., and Nausch, G. (2000). Hydrographic and current measurements in the area of the Angola–Benguela Front. *Journal of Physical Oceanography*, 30(10), 2589-2609.

Lebreton, B., Richard, P., Parlier, E.P., Guillou, G., and Blanchard, G.F. (2011). Trophic ecology of mullets during their spring migration in a European saltmarsh: a stable isotope study. *Estuarine, Coastal and Shelf Science*, 91(4), 502-510.

Lecomte, F., Grant, W.S., Dodson, J.J., Rodriquez-Sanchez, R., and Bowen, B.W. (2004). Living with uncertainty: genetic imprints of climate shifts in East Pacific anchovy (Engraulis mordax) and sardine (Sardinops sagax). *Molecular Ecology*, 13(8), 2169-2182.

Lee, S.C., Chang, J.T., and Tsu, Y.Y. (1995). Genetic relationships of four Taiwan mullets (Pisces: Perciformes: Mugilidae). *Journal of fish biology*, 46(1), 159-162.

Lefebure, T., Douady, C.J., Gouy, M., and Gibiert, J. (2006). Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution*, 40: 435–447.

Lemaire, C., Versini, J.J., and Bonhomme, F. (2005). Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. *Journal of Evolutionary Biology*, 18(1), 70-80.

Lemasson, L., and Rebert, J.P. (1973). Marine currents in the Ivorian Gulf. *Cah. ORSTOM, Ser. Oceanogr* , 11 , 67-95.

Lessios, H.A., and Robertson, D.R. (2006). Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings of the Royal Society of London B: Biological Sciences*, 273(1598), 2201-2208.

Lessios, H.A., Kane, J., and Robertson, D.R. (2003). Phylogeography of the pantropical sea urchin Tripneustes: contrasting patterns of population structure between oceans. *Evolution*, 57(9), 2026-2036.

Lessios, H.A., Kessing, B.D., and Pearse, J.S. (2001). Population structure and speciation in tropical seas: global phylogeography of the sea urchin Diadema. *Evolution*, 55(5), 955-975.

Lester, S.E., Ruttenberg, B.I., Gaines, S.D., and Kinlan, B.P. (2007). The relationship between dispersal ability and geographic range size. *Ecology Letters,* 10, 745-758.

Lévêque, C. (1997). *Biodiversity dynamics and conservation: the freshwater fish of tropical Africa*. Cambridge University Press.

Li, M.D. (1992). Biology and aquaculture of Mullet. *Nankai University*.

Li, S., and Chou, H.H. (2004). LUCY2: an interactive DNA sequence quality trimming and vector removal tool. *Bioinformatics*, 20(16), 2865-2866.

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25(11), 1451-1452.

Lietzke, C.E., Deser, C., and Vonder Haar, T.H. (2001). Evolutionary structure of the eastern Pacific double ITCZ based on satellite moisture profile retrievals. *Journal of climate*, 14(5), 743-751.

Lima, F.P., and Wethey, D.S. (2012). Three decades of high-resolution coastal sea surface temperatures reveal more than warming. *Nature communications*, 3, 704.

Lindberg, D.R. (1991). Marine biotic interchange between the northern and southern hemispheres. *Paleobiology*, 17(3), 308-324.

Linnaeus, C. (1758). Systema Naturae. Edition 10 (1). Holmiae. [=Stockholm] [Reprint Leipzig, 1894].

Liu, J.X., Gao, T.X., Zhuang, Z.M., Jin, X.S., Yokogawa, K., and Zhang, Y.P. (2006). Late Pleistocene divergence and subsequent population expansion of two closely related fish species, Japanese anchovy (*Engraulis japonicus*) and Australian anchovy (*Engraulis australis*). *Molecular Phylogenetics and Evolution*, 40(3), 712-723.

Liu, J.Y., Brown, C.L., and Yang, T.B. (2009). Population genetic structure and historical demography of grey mullet, Mugil cephalus, along the coast of China, inferred by analysis of the mitochondrial control region. *Biochemical Systematics and Ecology*, 37(5), 556-566.

Liu, J.Y., Brown, C.L., and Yang, T.B. (2010). Phylogenetic relationships of mullets (Mugilidae) in China Seas based on partial sequences of two mitochondrial genes. *Biochemical Systematics and Ecology*, 38(4), 647-655.

Livi, S., Sola, L., and Crosetti, D. (2011). Phylogeographic relationships among worldwide populations of the cosmopolitan marine species, the striped grey mullet (Mugil cephalus), investigated by partial cytochrome b gene sequences. *Biochemical Systematics and Ecology*, 39(2), 121-131.

Llave, E., Hernández-Molina, F.J., Stow, D.A., Fernández-Puga, M.C., García, M., Vázquez, J.T., ... and Del Rio, V.D. (2007). Reconstructions of the Mediterranean Outflow Water during the quaternary based on the study of changes in buried mounded drift stacking pattern in the Gulf of Cadiz. *Marine Geophysical Researches*, 28(4), 379-394.

Longhurst, A.R. (2010). *Ecological geography of the sea*. Elsevier.

Longhurst, A.R. (1962). A review of the oceanography of the Gulf of Guinea. *Bull. Inst. Fr. Afr. Noire*, 24, 633–663.

Lough, J.M. (2012). Small change, big difference: sea surface temperature distributions for tropical coral reef ecosystems, 1950–2011. *Journal of Geophysical Research: Oceans*, 117(C9).

Lowry, D.B., Hoban, S., Kelley, J.L., Lotterhos, K.E., Reed, L.K., Antolin, M.F., and Storfer, A. (2017). Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular ecology resources*, 17(2), 142-152.

Ludt, W.B. (2018). Determining the Drivers of Anti-Tropical Distributions Across the Fish Tree of Life. *LSU Doctoral Dissertations*. 4649.

Ludt, W.B., Rocha, L.A., Erdmann, M.V., and Chakrabarty, P. (2015). Skipping across the tropics: The evolutionary history of sawtail surgeonfishes (Acanthuridae: Prionurus). *Molecular phylogenetics and evolution*, 84, 166-172.

Lutjeharms, J.R.E., and Meeuwis, J.M. (1987). The extent and variability of South-East Atlantic

upwelling. *South African Journal of Marine Science*, 5(1), 51-62.

Mabuchi, K., Miya, M., Azuma, Y., and Nishida, M. (2007). Independent evolution of the specialized pharyngeal jaw apparatus in cichlid and labrid fishes. *BMC Evolutionary Biology*, 7(1), 10.

MacKenzie, K., Campbell, N., Mattiucci, S., Ramos, P., Pinto, A.L., and Abaunza, P. (2008). Parasites as biological tags for stock identification of Atlantic horse mackerel *Trachurus trachurus* L. *Fisheries Research*, 89(2), 136-145.

Magdy, M., Eshak, M.G., and Rashed, M.A.S. (2016). Genetic structure of Mugil cephalus L. populations from the northern coast of Egypt. *Veterinary world*, 9(1), 53.

Maggs, C.A., Castilho, R., Foltz, D., Henzler, C., Jolly, M.T., Kelly, J., ... and Viard, F. (2008). Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. *Ecology*, 89(sp11), S108-S122.

Magoulas, A., Tsimenides, N., and Zouros, E. (1996). Mitochondrial DNA phylogeny and the reconstruction of the population history of a species: the case of the European anchovy (Engraulis encrasicolus). *Molecular Biology and Evolution*, 13(1), 178-190.

Maillard, C., and Käse, R. (1989). The near-surface flow in the subtropical gyre south of the Azores. *Journal of Geophysical Research: Oceans*, 94(C11), 16133-16140.

Malanotte-Rizzoli, P., and Bergamasco, A. (1989). The circulation of the eastern mediterranean. Part 1. *Oceanologica Acta*, 12(4), 335-351.

Maley, J. (1991). The African rain forest vegetation and palaeoenvironments during late Quaternary. *Climatic change*, 19(1-2), 79-98.

Mamoozadeh, N.R., Graves, J.E., & McDowell, J.R. (2019). Genome-wide SNPs resolve spatiotemporal patterns of connectivity within striped marlin (Kajikia audax), a broadly distributed and highly migratory pelagic species. *Evolutionary Applications*. 1752-4571

Mann, B.Q. (2000). South African linefish status reports. *Oceanographic Research Institute Special Publication*, 7: 1–260.

Marko, P.B. (2002). Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology and Evolution*, 19(11), 2005-2021.

Marlow, J.R., Lange, C.B., Wefer, G., and Rosell-Melé, A. (2000). Upwelling intensification as part of the Pliocene-Pleistocene climate transition. *Science*, 290(5500), 2288-2291.

Masello, J.F., Quillfeldt, P., Munimanda, G.K., Klauke, N., Segelbacher, G., Schaefer, H.M., ... and Moodley, Y. (2011). The high Andes, gene flow and a stable hybrid zone shape the genetic structure of a wide-ranging South American parrot. *Frontiers in Zoology*, 8(1), 16.

Maslin, M.A., and Brierley, C.M. (2015). The role of orbital forcing in the Early Middle Pleistocene Transition. *Quaternary International*, 389, 47-55.

Maslin, M.A., Li, X.S., Loutre, M.F., and Berger, A. (1998). The contribution of orbital forcing to the progressive intensification of Northern Hemisphere glaciation. *Quaternary Science Reviews*, 17(4-5), 411-426.

Mayr, E. (1942). *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard University Press.

Mayr, E. (1954). Geographic speciation in tropical echinoids. *Evolution*, 8(1), 1-18.

Mayr, G. (2013). The age of the crown group of passerine birds and its evolutionary significance–molecular calibrations versus the fossil record. *Systematics and Biodiversity*, 11(1), 7-13.

McClellan, A.J., Tam, S., Kaganovich, D., and Frydman, J. (2005). Protein quality control: chaperones culling corrupt conformations. *Nature cell biology*, 7(8), 736.

McKeown, N.J., Arkhipkin, A.I., and Shaw, P.W. (2015). Integrating genetic and otolith microchemistry data to understand population structure in the Patagonian Hoki (Macruronus magellanicus). *Fisheries Research*, 164: 1–7.

McKeown, N.J., Arkhipkin, A.I., and Shaw, P.W. (2016). Regional genetic population structure and fine scale genetic cohesion in the Southern blue whiting Micromesistius australis. *Fisheries Research*, 185: 176–184.

McKeown, N.J., Arkhipkin, A.I., and Shaw, P.W. (2019). Genetic analysis reveals historical and contemporary population dynamics in the longfin squid *Doryteuthis gahi*: implications for cephalopod management and conservation. *ICES Journal of Marine Science*.

McKeown, N.J., Hauser, L., and Shaw, P.W. (2017). Microsatellite genotyping of brown crab Cancer pagurus

reveals fine scale selection and 'non-chaotic' genetic patchiness within a high gene flow system. *Marine Ecology Progress Series*, 566: 91–103.

McKeown, N.J., Robin, J.P., and Shaw, P.W. (2015). Species-specific PCR-RFLP for identification of early life history stages of squid and other applications to fisheries research. *Fisheries research*, 16: 207–209.

McLaverty, K. (2012). A re-evaluation of the life history strategy of Cape horse mackerel, *Trachurus capensis* in the southern Benguela (Doctoral dissertation, The University of Cape Town). Available from Open UCT library (https://open.uct.ac.za/handle/11427/12101).

McMahan, C.D., Davis, M.P., Domínguez-Domínguez, O., García-de-León, F.J., Doadrio, I., and Piller, K.R. (2013). From the mountains to the sea: phylogeography and cryptic diversity within the mountain mullet, Agonostomus monticola (Teleostei: Mugilidae). *Journal of Biogeography*, 40(5), 894-904.

Mcmillen-Jackson, A.L., and Bert, T.M. (2003). Disparate patterns of population genetic structure and population history in two sympatric penaeid shrimp species (Farfantepenaeus aztecus and Litopenaeus setiferus) in the eastern United States. *Molecular Ecology*, 12(11), 2895-2905.

McMillen-Jackson, A.L., and Bert, T.M. (2004). Genetic diversity in the mtDNA control region and population structure in the pink shrimp Farfantepenaeus duorarum. *Journal of Crustacean Biology*, 24(1), 101-109.

Meeuwis, J.M., and Lutjeharms, J.R.E. (1990). Surface thermal characteristics of the Angola-Benguela front. *South African Journal of Marine Science*, 9(1), 261-279.

Mellin, C., Mouillot, D., Kulbicki, M., Mcclanahan, T.R., Vigliola, L., Bradshaw, C.J.A., ...and Friedlander, A.M. (2016). Humans and seasonal climate variability threaten large-bodied coral reef fish with small ranges. *Nature Communications*, 7, 10491.

Mello, B., Tao, Q., Tamura, K., and Kumar, S. (2016). Fast and accurate estimates of divergence times from big data. *Molecular biology and evolution*, 34(1), 45-50.

Menezes, M.R., Martins, M., and Naik, S. (1992). Interspecific genetic divergence in grey mullets from the Goa region. *Aquaculture*, 105(2), 117-129.

Mensah, M.A., and Quaatey, S.N.K. (2002). 17 An overview of the fishery resources and fishery research in the Gulf of Guinea. In *Large Marine Ecosystems* (Vol. 11, pp. 227-iii). Elsevier.

Meyer, A. (1994). DNA technology and phylogeny of fish. In Genetics and Evolution of Aquatic Organisms (Beaumont, A.R., ed.), pp. 219–249. London: Chapman and Hall.

Meyer, C.P., and Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology,* 3: e422.

Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., Carvalho, G.R., ... and Hansen, J.H. (2014). Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (Merluccius merluccius). *Molecular Ecology*, 23(1), 118-135.

Millot, C. (1999). Circulation in the western Mediterranean Sea. *Journal of Marine Systems*, 20(1-4), 423-442.

Millot, C., and Taupier-Letage, I. (2005). Circulation in the Mediterranean sea. In *The Mediterranean Sea* (pp. 29-66). Springer, Berlin, Heidelberg.

Mittelstaedt, E. (1991). The ocean boundary along the northwest African coast: Circulation and oceanographic properties at the sea surface. *Progress in Oceanography*, 26(4), 307-355.

Mix, A.C., Ruddiman, W.F., and McIntyre, A. (1986). Late Quaternary paleoceanography of the Tropical Atlantic, 1: Spatial variability of annual mean sea-surface temperatures, 0-20,000 years BP. *Paleoceanography and Paleoclimatology*, 1(1), 43-66.

Morales-Nin, B., & Moranta, J. (1997). Life history and fishery of the common dentex (Dentex dentex) in Mallorca (Balearic Islands, western Mediterranean). *Fisheries Research*, 30(1-2), 67-76.

Moreira, C., Correia, A.T., Vaz-Pires, P., and Froufe, E. (2019). Genetic diversity and population structure of the blue jack mackerel *Trachurus picturatus* across its western distribution. *Journal of fish biology*.

Moritz, C. (2002). Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic biology*, 51(2), 238-254.

Moroshkin, K.V., Bubnov, V.A., and Bulatov R.P. (1970). Water circulation in the eastern South Atlantic Ocean. *OceanoloRY*, 10(1), 27-34.

Mullins, R.B., McKeown, N.J., Sauer, W.H., and Shaw, P.W. (2018). Genomic analysis reveals multiple mismatches between biological and management units in yellowfin tuna (Thunnus albacares). *ICES Journal of Marine Science*, 75(6), 2145-2152.

Murgia, R., Tola, G., Archer, S.N., Vallerga, S., and Hirano, J. (2002). Genetic identification of grey mullet species (Mugilidae) by analysis of mitochondrial DNA sequence: application to identify the origin of pro-

cessed ovary products (bottarga). *Marine Biotechnology*, 4(2), 119-126.

Murta, A. G., Pinto, A. L., and Abaunza, P. (2008). Stock identification of horse mackerel (*Trachurus trachurus*) through the analysis of body shape. *Fisheries Research*, 89(2), 152-158.

Murta, A.G. (2000). Morphological variation of horse mackerel (*Trachurus trachurus*) in the Iberian and North African Atlantic: implications for stock identification. *ICES Journal of Marine Science,* 57(4), 1240-1248.

Muths, D., Grewe, P., Jean, C., and Bourjea, J. (2009). Genetic population structure of the Swordfish (Xiphias gladius) in the southwest Indian Ocean: Sex-biased differentiation, congruency between markers and its incidence in a way of stock assessment. *Fisheries Research,* 97, 263-269.

Muths, D., Le Couls, S., Evano, H., Grewe, P., and Bourjea, J. (2013). Multi-Genetic Marker Approach and Spatio-Temporal Analysis Suggest There Is a Single Panmictic Population of Swordfish Xiphias gladius in the Indian Ocean. *Plos One,* 8.

Muths, D., Tessier, E., and Bourjea, J. (2015). Genetic structure of the reef grouper *Epinephelus merra* in the West Indian Ocean appears congruent with biogeographic and oceanographic boundaries. *Marine Ecology-an Evolutionary Perspective*, 36: 447–461.

Muths, D., Tessier, E., Gouws, G., Craig, M., Mwale, M., Mwaluma, J., and Bourjea, J. (2011). Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the SW Indian Ocean. *Marine Ecology Progress Series*, 443: 167–180.

Nabholz, B., Glémin, S., and Galtier, N. (2007). Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. *Molecular biology and evolution*, 25(1), 120-130.

Nabholz, B., Glémin, S., and Galtier, N. (2009). The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evolutionary Biology*, 9 (1), 54.

Nabholz, B., Mauffrey, J.F., Bazin, E., Galtier, N., and Glemin, S. (2008). Determination of mitochondrial genetic diversity in mammals. *Genetics*, 178(1), 351-361.

Naciri, M., Lemaire, C., Borsa, P., and Bonhomme, F. (1999). Genetic study of the Atlantic/Mediterranean transition in sea bass (Dicentrarchus labrax). *Journal of Heredity*, 90(6), 591-596.

Naish, K.A. (1990). The stock identification of the Cape horse mackerel, *Trachurus trachurus capensis* (*Pisces: Carangidae*). Masters thesis, Rhodes University. Available from National Research Foundation NEXUS library (http://nrfnexus.nrf.ac.za/handle/20.500.11892/22864).

Naish, K.A., Hecht, T., and Payne, A.I.L. (1991). Growth of Cape horse mackerel *Trachurus trachurus capensis* off South Africa. *South African Journal of Marine Science*, 10(1), 29-35.

Nakamura, Y., Shibuno, T., Suzuki, N., Nakamori, J., Kanashiro, K., and Watanabe, Y. (2010). Interspecific variations in age and size at settlement of 8 emperor fishes (Lethrinidae) at the southern Ryukyu Islands, Japan. *Fisheries Science,* 76, 503-510.

Naranjo, S., Carballo, J.L., and Garcia-Gomez, J.C. (1998) Towards a knowledge of marine boundaries using ascidians as indicators: characterising transition zones for species distribution along Atlantic-Mediterranean shores. *Biological Journal of the Linnean Society*, 64, 151–177

Narayan, N., Paul, A., Mulitza, S., and Schulz, M. (2010). Trends in coastal upwelling intensity during the late 20th century. *Ocean Science*, 6(3), 815-823.

Narum, S.R., and Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular ecology resources*, 11, 184-194.

Navarro-Pérez, E., and Barton, E.D. (2001). Seasonal and interannual variability of the Canary Current. *Scientia Marina*, 65(S1), 205-213.

Nei, M. (1987). Molecular Evolutionary Genetics. Columbia Univ. Press, New York

Nekrasov, V.V. (1976). Systematic position, morphological characteristics and range of Carangidae of the Indian Ocean. Moscow, *Central fisheries Technical and Economic Information Research Institute*. Manuscript Department 6.

Nelson, J.S. (2006). Fishes of the world. *Hoboken: John Wiley and Sons*.

Nesis, K.N. (1993). Cephalopods from the Saya de Malha Bank, Indian Ocean. *Transactions of the PP Shirshov Institute of Oceanology*, 128: 26–39.

Nichols, R. (2001). Gene trees and species trees are not the same. *Trends in Ecology and Evolution*, 16(7), 358 -364.

Nielsen, E.E., Cariani, A., Mac Aoidh, E., Maes, G.E., Milano, I., Ogden, R., ... and Bekkevold, D. (2012). Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature communications*, 3, 851.

Nielsen, E.E., Hansen, M.M., and Loeschcke, V. (1997). Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison of genetic composition over 60 years. *Molecular Ecology*, 6(5), 487-492.

Niemiller, M.L., Fitzpatrick, B.M., and Miller, B.T. (2008). Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene genealogies. *Molecular ecology*, 17 (9), 2258-2275.

Nikolaev, S.D. (2006). Principal features of the World Ocean climate change in the Cenozoic and their possible causes. *Oceanology*, 46(4), 513-525.

Nilson, T. (1982). Alluvial fan deposits, in Scholle, P. A., and Spearing, P., eds., Sandstone depositional environments: American Association of Petroleum Creologists Memoir 31, p. 49-86.

Nohara, M., Nishida, M., Manthacitra, V., and Nishikawa, T. (2004). Ancient phylogenetic separation between Pacific and Atlantic cephalochordates as revealed by mitochondrial genome analysis. *Zoological science*, 21(2), 203-211.

Nolf, D. (2003). Fish otoliths from the Santonian of the Pyrenean faunal province, and an overview of all otolith-documented North Atlantic Late Cretaceous teleosts. *Bulletin-Institut royal des sciences naturelles de Belgique: Bulletin-Koninklijk Belgisch Institut voor Natuurwetenschappen: Aardwetenschappen. Sciences de la terre*, 73, 155.

Nolf, D., and Aguilera, O. (1998). Fish otoliths from the Cantaure Formation (early Miocene of Venezuela). *Bulletin de l'Institut royal des Sciences naturelles de Belgique, Sciences de la Terre*, 68, 237-262.

Noor, M.A., Grams, K.L., Bertucci, L.A., and Reiland, J. (2001). Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences*, 98(21), 12084-12088.

Norris, R.D. (1998). Miocene-Pliocene surface-water hydrography of the eastern equatorial Atlantic. In *proceedings-ocean drilling program scientific results* (pp. 539-556). National Science Foundation

Nosil, P., Crespi, B.J., and Sandoval, C.P. (2002). Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature*, 417(6887), 440.

Oberhänsli, H. (1991). Upwelling signals at the northeastern Walvis Ridge during the past 500,000 years. *Paleoceanography and Paleoclimatology*, 6(1), 53-71.

Okumura, Y., and Xie, S.P. (2004). Interaction of the Atlantic equatorial cold tongue and the African monsoon. *Journal of Climate*, 17(18), 3589-3602.

Orrell, T.M., Carpenter, K.E., Musick, J.A., & Graves, J.E. (2002). Phylogenetic and biogeographic analysis of the Sparidae (Perciformes: Percoidei) from cytochrome b sequences. *Copeia*, *2002*(3), 618-631.

Ott, A., Liu, S., Schnable, J.C., Yeh, C.T.E., Wang, K.S., and Schnable, P.S. (2017). tGBS® genotyping-by-sequencing enables reliable genotyping of heterozygous loci. *Nucleic acids research*, 45(21), e178-e178.

Page, T.J., and Hughes, J.M. (2010). Comparing the performance of multiple mitochondrial genes in the analysis of Australian freshwater fishes. *Journal of Fish Biology*, *77*(9), 2093-2122.

Palero, F., Lopes, J., Abelló, P., Macpherson, E., Pascual, M., and Beaumont, M.A. (2009). Rapid radiation in spiny lobsters (Palinurus spp) as revealed by classic and ABC methods using mtDNA and microsatellite data. *BMC Evolutionary Biology*, 9(1), 263.

Palumbi, S.R. (1992). Marine speciation on a small planet. *Trends in Ecology and Evolution*, *7*(4), 114-118.

Palumbi, S.R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annual review of ecology and systematics*, 25(1), 547-572.

Palumbi, S.R., Cipriano, F., and Hare, M.P. (2001). Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. *Evolution*, 55(5), 859-868.

Panhuis, T.M., Butlin, R., Zuk, M., and Tregenza, T. (2001). Sexual selection and speciation. *Trends in ecology and evolution*, 16(7), 364-371.

Pankhurst, N.W., & Munday, P.L. (2011). Effects of climate change on fish reproduction and early life history stages. *Marine and Freshwater Research*, 62(9), 1015-1026.

Panova, M., Hollander, J., & Johannesson, K. (2006). Site-specific genetic divergence in parallel hybrid zones suggests nonallopatric evolution of reproductive barriers. *Molecular Ecology*, 15(13), 4021-4031.

Papasotiropoulos, V., Klossa-Kilia, E., Alahiotis, S.N., and Kilias, G. (2007). Molecular phylogeny of grey mullets (Teleostei: Mugilidae) in Greece: evidence from sequence analysis of mtDNA segments. *Biochemical genetics*, 45(7-8), 623.

Papasotiropoulos, V., Klossa-Kilia, E., Kilias, G., and Alahiotis, S. (2002). Genetic divergence and phylogenetic relationships in grey mullets (Teleostei: Mugilidae) based on PCR–RFLP analysis of mtDNA segments. *Biochemical Genetics*, 40(3-4), 71-86.

Pardo-Gandarillas, M.C., Ibáñez, C.M., Yamashiro, C., Méndez, M.A., and Poulin, E. (2018). Demographic inference and genetic diversity of *Octopus mimus (Cephalopoda: Octopodidae)* throughout the Humboldt Current System. *Hydrobiologia*, 808(1), 125-135.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.*, 37, 637-669.

Parry, M. (2006). Feeding behavior of two ommastrephid squids Ommastrephes bartramii and Sthenoteuthis oualaniensis off Hawaii. *Marine Ecology Progress Series*, 318, 229-235.

Patarnello, T., Volckaert, F.A., and Castilho, R. (2007). Pillars of Hercules: is the Atlantic–Mediterranean transition a phylogeographical break?. *Molecular ecology*, 16(21), 4426-4444.

Payne, A.I. (1976). The relative abundance and feeding habits of the grey mullet species occurring in an estuary in Sierra Leone, West Africa. *Marine Biology*, 35(3), 277-286.

Peakall, R., and Smouse, P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, 28**:** 2537–2539.

Peakall, R.O.D., and Smouse, P.E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular ecology note*s, 6(1), 288-295.

Pecoraro, C., Babbucci, M., Villamor, A., Franch, R., Papetti, C., Leroy, B., ... and Murua, H. (2016). Methodological assessment of 2b-RAD genotyping technique for population structure inferences in yellowfin tuna (Thunnus albacares). *Marine genomics*, 25, 43-48.

Pecoraro, C., Zudaire, I., Bodin, N., Murua, H., Taconet, P., Diaz-Jaimes, P., ... and Chassot, E. (2017). Putting all the pieces together: integrating current knowledge of the biology, ecology, fisheries status, stock structure and management of yellowfin tuna (Thunnus albacares). *Reviews in fish biology and fisheries*, 27(4), 811-841.

Peijnenburg, K.T., Breeuwer, J.A., Pierrot-Bults, A.C., and Menken, S.B. (2004). Phylogeography of the planktonic chaetognath Sagitta setosa reveals isolation in European seas. *Evolution*, 58(7), 1472-1487.

Pelegrí, J.L., Peña-Izquierdo, J. (2015). Eastern boundary currents off north-west Africa. In: Valdés L, Déniz-González I (eds), Oceanographic and biological features in the Canary Current Large Marine Ecosystem. IOC Technical Series No. 115. Paris: IOC-UNESCO. pp 81–92

Penzo, E., Gandolfi, G., Bargelloni, L., Colombo, L., and Patarnello, T. (1998). Messinian salinity crisis and the origin of freshwater lifestyle in western Mediterranean gobies. *Molecular Biology and Evolution*, 15 (11), 1472-1480.

Pérez-Losada, M., Browne, E.B., Madsen, A., Wirth, T., Viscidi, R.P., and Crandall, K.A. (2006). Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data. *Infection, Genetics and Evolution*, 6(2), 97-112.

Pérez-Losada, M., Guerra, Á., and Sanjuan, A. (1999). Allozyme differentiation in the cuttlefish Sepia officinalis (Mollusca: Cephalopoda) from the NE Atlantic and Mediterranean. *Heredity*, 83(3), 280.

Pérez-Losada, M., Guerra, A., Carvalho, G.R., Sanjuan, A., and Shaw, P.W. (2002). Extensive population subdivision of the cuttlefish Sepia officinalis (Mollusca: Cephalopoda) around the Iberian Peninsula indicated by microsatellite DNA variation. *Heredity*, 89(6), 417.

Pérez-Losada, M., Nolte, M.J., Crandall, K.A., and Shaw, P.W. (2007). Testing hypotheses of population structuring in the Northeast Atlantic Ocean and Mediterranean Sea using the common cuttlefish Sepia officinalis. *Molecular Ecology*, 16(13), 2667-2679.

Peterson, R.G., and Stramma, L. (1991). Upper-level circulation in the South Atlantic Ocean. *Progress in oceanography*, 26(1), 1-73.

Petren, K., Grant, P.R., Grant, B.R., and Keller, L.F. (2005). Comparative landscape genetics and the adaptive radiation of Darwin's finches: the role of peripheral isolation. *Molecular Ecology*, 14(10), 2943-2957.

Phillips, S.J., Anderson, R.P., and Schapire, R.E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* 190, 231-259.

Pierce GJ, Bailey N, Stratoudakis Y, and Newton A. (1998). Distribution and abundance of the fished population of *Loligo forbesi* in Scottish waters: analysis of research cruise data. *ICES Journal of Marine Science: Journal du Conseil* 55: 14–33.

Piñeira, J., Quesada, H., Rolán-Alvarez, E., and Caballero, A. (2008). Genetic discontinuity associated with an environmentally induced barrier to gene exchange in the marine snail Littorina saxatilis. *Marine Ecology Progress Series*, 357, 175-184.

Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L., and Estoup, A. (2004). GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of heredity*, 95(6), 536-539.

Pisias, N.G., and Moore Jr, T.C. (1981). The evolution of Pleistocene climate: a time series approach. *Earth and Planetary Science Letters*, 52(2), 450-458.

Pisias, N.G., Mayer, L.A., and Mix, A.C. (1995). Paleoceanography of the eastern equatorial Pacific during the Neogene: synthesis of Leg 138 drilling results. Proc. ODP, Sci. Res. 138, 5–21.

Pitcher, G.C., and Weeks, S.J. (2006). 7 The variability and potential for prediction of harmful algal blooms in the southern Benguela ecosystem. In *Large marine ecosystems* (Vol. 14, pp. 125-146). Elsevier.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., ... and Vogler, A.P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic biology*, 55(4), 595-609.

Pörtner, H.O., and Knust, R. (2007). Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *science*, 315(5808), 95-97.

Pörtner, H.O., Berdal, B., Blust, R., Brix, O., Colosimo, A., De Wachter, B., ... & Lannig, G. (2001). Climate induced temperature effects on growth performance, fecundity and recruitment in marine fish: developing a hypothesis for cause and effect relationships in Atlantic cod (Gadus morhua) and common eelpout (Zoarces viviparus). *Continental Shelf Research*, 21(18-19), 1975-1997.

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Molecular biology and evolution*, 25(7): 1253–1256.

Posada, D., and Crandall, K.A. (1998). Modeltest: testing the model of DNAsubstitution. *Bioinformatics (Oxford, England)*, 14(9), 817-818.

Potts, W.M., Booth, A.J., Richardson, T.J., and Sauer, W.H. (2014b). Ocean warming affects the distribution and abundance of resident fishes by changing their reproductive scope. *Reviews in fish biology and fisheries*, 24(2), 493-504.

Potts, W.M., Henriques, R., Santos, C.V., Munnik, K., Ansorge, I., Dufois, F., ... and Shaw, P.W. (2014a). Ocean warming, a rapid distributional shift, and the hybridization of a coastal fish species. *Global Change Biology*, 20(9), 2765-2777.

Prell, W.L., Gardner, J.V., Bé, A.W., and Hays, J.D. (1976). Equatorial Atlantic and Caribbean foraminiferal assemblages, temperatures and circulation: interglacial and glacial comparisons. *Geol. Soc. America Mem*, 145, 247-266.

Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945-959.

Puechmaille, S.J. (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, 16(3), 608-627.

Pujolar, J.M., Roldan, M.I., and Pla, C. (2002). A genetic assessment of the population structure of swordfish (Xiphias gladius) in the Mediterranean Sea. *Journal of Experimental Marine Biology and Ecology*, 276(1-2), 19-29.

Queller, D.C., and Goodnight, K.F. (1989). Estimating relatedness using genetic markers. *Evolution*, 258-275.

Quesada, H., Beynon, C.M., and Skibinski, D.O. (1995). A mitochondrial DNA discontinuity in the mussel Mytilus galloprovincialis Lmk: pleistocene vicariance biogeography and secondary intergradation. *Molecular Biology and Evolution*, 12(3), 521-524.

Quinteiro, J., Manent, P., Pérez-Diéguez, L., González, J.A., Almeida, C., Lopes, E., ... and González-Henríquez, N. (2015). Phylogeography of a marine insular endemic in the Atlantic Macaronesia: the Azorean barnacle, Megabalanus azoricus (Pilsbry, 1916). *PloS one*, 10(4), e0124707.

Quinteiro, J., Rodríguez-Castro, J., and Rey-Méndez, M. (2007). Population genetic structure of the stalked barnacle Pollicipes pollicipes (Gmelin, 1789) in the northeastern Atlantic: influence of coastal currents

and mesoscale hydrographic structures. *Marine Biology*, 153(1), 47-60.

Quoy, J.R., and Gaimard, J.P. (1832). Mollusques. In: Voyage de découvertes de l'Astrolabe pendant les années 1826–1827–1828–1829. *Zoologie*, 2: 1–320.

Ragionieri, L., Cannicci, S., Schubart, C.D., and Fratini, S. (2010). Gene flow and demographic history of the mangrove crab Neosarmatium meinerti: A case study from the western Indian Ocean. *Estuarine, Coastal and Shelf Science*, 86(2), 179-188.

Ragionieri, L., Fratini, S., Vannini, M., and Schubart, C.D. (2009). Phylogenetic and morphometric differentiation reveal geographic radiation and pseudo-cryptic speciation in a mangrove crab from the Indo-West Pacific. *Molecular Phylogenetics and Evolution*, 52, 825-834.

Rambaut, A., and Bromham, L. (1998). Estimating divergence dates from molecular sequences. *Molecular biology and evolution*, 15(4), 442-448.

Ramos, J.E., Pecl, G.T., Moltschaniwskyj, N.A., Strugnell, J.M., León, R.I., and Semmens, J.M. (2014). Body size, growth and life span: implications for the polewards range shift of Octopus tetricus in south-eastern Australia. *PLoS One*, 9(8), e103480.

Rand, D.M. (2001). The units of selection on mitochondrial DNA. *Annual Review of Ecology and Systematics*, 32(1), 415-448.

Randi, E., and Lucchini, V. (1998). Organization and evolution of the mitochondrial DNA control region in the avian genus Alectoris. *Journal of Molecular Evolution*, *47*(4), 449-462.

Rang, M.S. (1837). Documents pour servir à l'histoire naturelle des Céphalopodes crypyobranches. *Magasin de Zoologie, Mollusques, classe V* 7: 1–77.

Raymond, M., and Rousset, F. (1995). An exact test for population differentiation. *Evolution*, 49(6), 1280-1283.

Raymond, M., and Rousset, F. (1995). GENEPOP (Version-1.2) – population-genetics software for exact tests and ecumenicism. *Journal of Heredity,* 86, 248-249.

Reichenbacher, B., and Weidmann, M. (1992). Fisch-otolithen aus der Oligo/Mioz¨anen molasse der West-Schweiz und der Haute Savoie (Frankreich). Stuttgarter Beitr¨age zur Naturkunde, Serie B (Geologie und Palaontologie), 184, 1–83.

Reichow, D., and Smith, M.J. (2001). Microsatellites reveal high levels of gene flow among populations of the California squid Loligo opalescens. *Molecular Ecology*, 10(5), 1101-1109.

Reid, A., Jereb, P., and Roper, C.F.E. (2005). Family Sepiidae. In: Jereb, P., Roper, C.F.E. (eds), *Cephalopods of the world. An annotated and illustrated catalogue of species known to date. Vol. 1. Chambered Nautiluses and Sepioids (Nautilidae, Sepiidae, Sepiolidae, Sepiadariidae, Idiosepiidae and Spirulidae)*. *FAO Species Catalogue for Fisheries Purposes,* No. 4. Vol. 1. Rome: FAO. pp 57–152.

Reid, K., Hoareau, T.B., Graves, J.E., Potts, W.M., dos Santos, S.M.R., Klopper, A.W., and Bloomer, P. (2016). Secondary contact and asymmetrical gene flow in a cosmopolitan marine fish across the Benguela upwelling zone. *Heredity*, 117(5), 307.

Reiss, H., Hoarau, G., Dickey-Collas, M., and Wolff, W.J. (2009). Genetic population structure of marine fish: mismatch between biological and fisheries management units. *Fish and Fisheries*, 10(4), 361-395.

Rey, O.T.(1996). Estratigrafía de la península de Paraguaná, Venezuela. *Revista de la Facultad de Ingeniería, Universidad Central de Venezuela*, 11(1), 35-45.

Rhein, M., Rintoul, S.R., Aoki, S., Campos, E., Chambers, D., Feely, R.A., ... and Mauritzen, C. (2013). Observations: ocean Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change ed TF Stocker *et al.*. *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*.

Richards, E., Servedio, M., and Martin, C. (2018). New criteria for sympatric speciation in the genomic era. *bioRxiv*, 367623.

Richardson, P.L., and Philander, S.G.H. (1987). The seasonal variations of surface currents in the tropical Atlantic Ocean: A comparison of ship drift data with results from a general circulation model. *Journal of Geophysical Research: Oceans*, 92(C1), 715-724.

Ridgway, T., and Sampayo, E.M. (2005). Population genetic status of the Western Indian Ocean: What do we know?. *Western Indian Ocean Journal of Marine Science*, 4(1), 1-10.

Ridley, M. 2004 Evolution. Oxford, UK: Blackwell Publishing.

Rieseberg, L.H., Kim, S.C., Randell, R.A., Whitney, K.D., Gross, B.L., Lexer, C., and Clay, K. (2007). Hybridization

and the colonization of novel habitats by annual sunflowers. *Genetica*, 129(2), 149-165.

Roberts, C.M. (1997). Connectivity and management of Caribbean coral reefs. *Science*, *278*(5342), 1454-1457.

Robinson, A.R., Leslie, W.G., Theocharis, A., and Lascaratos, A. (2001). Mediterranean sea circulation. *Ocean currents: a derivative of the Encyclopedia of Ocean Sciences*, 1689-1705.

Robinson, R.S., Meyers, P.A., and Murray, R.W. (2002). Geochemical evidence for variations in delivery and deposition of sediment in Pleistocene light–dark color cycles under the Benguela Current Upwelling System. *Marine Geology*, 180(1), 249-270.

Rocha-Olivares, A., Garber, N.M., and Stuck, K.C. (2000). High genetic diversity, large inter-oceanic divergence and historical demography of the striped mullet. *Journal of Fish Biology*, 57(5), 1134-1149.

Rodhouse, P.G., Pierce, G.J., Nichols, O.C., Sauer, W.H., Arkhipkin, A.I., and Laptikhovsky, V.V. (2014). Environmental effects on cephalopod population dynamics: implications for management of fisheries. *Advances in Marine Biology*, 67: 99–233.

Roeleveld, M.A. (1972). A review of the Sepiidae (Cephalopoda) of Southern Africa. PhD thesis, Stellenbosch University, South Africa.

Roeleveld, M.A.C. (1998). The status and importance of cephalopod systematics in southern Africa. *South African Journal of Marine Science*, 20: 1–16.

Rogers, A.R., and Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular biology and evolution*, 9(3), 552-569.

Rogers, D.S., and González, M.W. (2010). Phylogenetic relationships among spiny pocket mice (Heteromys) inferred from mitochondrial and nuclear sequence data. *Journal of Mammalogy*, 91(4), 914-930.

Rogerson, M., Rohling, E.J., Bigg, G.R., and Ramirez, J. (2012). Paleoceanography of the Atlantic-Mediterranean exchange: Overview and first quantitative assessment of climatic forcing. *Reviews of Geophysics*, 50(2).

Rögl, F. (1999). Mediterranean and Paratethys. Facts and hypotheses of an Oligocene to Miocene paleogeography (short overview). *Geologica carpathica*, 50(4), 339-349.

Rohling, E.J., Foster, G.L., Grant, K.M., Marino, G., Roberts, A.P., Tamisiea, M.E., and Williams, F. (2014). Sea-level and deep-sea-temperature variability over the past 5.3 million years. *Nature*, 508(7497), 477.

Ronquist, F., and Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12), 1572-1574.

Rosenblatt, R.H., and Waples, R.S. (1986). A genetic comparison of allopatric populations of shore fish species from the eastern and central Pacific Ocean: dispersal or vicariance?. *Copeia*, 275-284.

Rossi, A.R., Capula, M., Crosetti, D., Campton, D.E., and Sola, L. (1998a). Genetic divergence and phylogenetic inferences in five species of Mugilidae (Pisces: Perciformes). *Marine Biology*, 131(2), 213-218.

Rossi, A.R., Capula, M., Crosetti, D., Sola, L., and Campton, D.E. (1998b). Allozyme variation in global populations of striped mullet, Mugil cephalus (Pisces: Mugilidae). *Marine Biology*, 131(2), 203-212.

Rossi, A.R., Gornung, E., and Crosetti, D. (1997). Cytogenetic analysis of Liza ramada (Pisces, Perciformes) by different staining techniques and fluorescent in situ hybridization. *Heredity*, 79(1), 83.

Rossi, A.R., Gornung, E., Crosetti, D., De Innocentiis, S., and Sola, L. (2000). Cytogenetic analysis of Oedalechilus labeo (Pisces: Mugilidae), with a report of NOR variability. *Marine Biology*, 136(1), 159-162.

Rossi, A.R., Ungaro, A., De Innocentiis, S., Crosetti, D., and Sola, L. (2004). Phylogenetic analysis of Mediterranean Mugilids by allozymes and 16S mt-rRNA genes investigation: are the Mediterranean species of Liza monophyletic?. *Biochemical genetics*, 42(9-10), 301-315.

Rouault, M., Pohl, B., and Penven, P. (2010). Coastal oceanic climate change and variability from 1982 to 2009 around South Africa. *African Journal of Marine Science*, 32(2), 237-246.

Rousset, F. (2008). GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, 8: 103–106.

Roux, C., Fraisse, C., Romiguier, J., Anciaux, Y., Galtier, N., and Bierne, N. (2016). Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS biology*, 14(12), e2000234.

Royer, J. (2002). Modélisation des stocks de céphalopodes de Manche. *Thése de 3 éme cycle. Université de Caen: 242 p.*

Rueda, F.M., & Martínez, F.M. (2001). A review on the biology and potential aquaculture of Dentex dentex. *Reviews in Fish Biology and Fisheries*, 11(1), 57-70.

Ruiz-Cooley, R.I., Ballance, L.T., and McCarthy, M.D. (2013). Range expansion of the jumbo squid in the NE Pacific: δ15N decrypts multiple origins, migration and habitat use. *PLoS One*, 8(3), e59651.

Rundle, H.D., and Nosil, P. (2005). Ecological speciation. *Ecology letters*, 8(3), 336-352.

Rundle, H.D., and Schluter, D. (2004). Natural selection and ecological speciation in sticklebacks. *Adaptive speciation*, 19(3), 192-209.

Rutschmann, F. (2006). Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. *Diversity and Distributions*, 12(1), 35-48.

Ruzzante, D.E., Mariani, S., Bekkevold, D., Andre, C., Mosegaard, H., Clausen, L.A.W., . . . and Carvalho, G.R. (2006). Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring. *Proceedings of the Royal Society B-Biological Sciences,* 273(1593), 1459-1464.

Ruzzante, D.E., Walde, S.J., Macchi, P.J., Alonso, M., and Barriga, J.P. (2011). Phylogeography and phenotypic diversification in the Patagonian fish Percichthys trucha: the roles of Quaternary glacial cycles and natural selection. *Biological Journal of the Linnean Society*, 103(2), 514-529.

Ryan, W.B. (1973). Geodynamic implications of the Messinian crisis of salinity. In *Messinian events in the Mediterranean* (pp. 26-38). North-Holland Amsterdam.

Ryman, N., and Palm, S. (2006). POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes,* 6, 600-602.

Ryman, N., Utter, F., and Laikre, L. (1995). Protection of intraspecific biodiversity of exploited fishes. *Reviews in Fish Biology and Fisheries*, 5(4): 417–446.

Saccone, C., Attimonelli, M., and Sbisa, E. (1987). Structural elements highly preserved during the evolution of the D-loop-containing region in vertebrate mitochondrial DNA. *Journal of molecular evolution*, 26 (3), 205-211.

Saccone, C., Pesole, G., and Sbisa, E. (1991). The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *Journal of Molecular Evolution*, 33(1), 83-91.

Sakko, A.L. (1998). The influence of the Benguela upwelling system on Namibia's marine biodiversity. *Biodiversity and conservation*, 7(4), 419-433.

Sala-Bozano, M., Ketmaier, V., & Mariani, S. (2009). Contrasting signals from multiple markers illuminate population connectivity in a marine fish. *Molecular Ecology*, 18(23), 4811-4826.

Sala-Bozano, M., Mariani, S., Barratt, C.D., Sacchi, C., Boufana, B., and Coscia, I. (2015). Spatio-temporal variability in the population structure in North-east Atlantic stocks of horse mackerel (*Trachurus trachurus*). *Biology and Environment: Proceedings of the Royal Irish Academy,* 115(3), 211-220).

Salat, J. (1996). Review of hydrographic environmental factors that may influence anchovy habitats in northwestern Mediterranean. *Scientia Marina*, 60, 21–32.

Sanchez, G., Tomano, S., Yamashiro, C., Fujita, R., Wakabayashi, T., Sakai, M., and Umino, T. (2016). Population genetics of the jumbo squid Dosidicus gigas (Cephalopoda: Ommastrephidae) in the northern Humboldt Current system based on mitochondrial and microsatellite DNA markers. *Fisheries research*, 175, 1-9.

Sang, T.K., Chang, H.Y., Chen, C.T., & Hui, C.F. (1994). Population structure of the Japanese eel, Anguilla japonica. *Molecular Biology and Evolution*, 11(2), 250-260.

Sanderson, M.J. (2003). r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, 19(2), 301-302.

Santini, F., May, M.R., Carnevale, G., and Moore, B.R. (2015). Bayesian Inference of Divergence Times and Feeding Evolution in Grey Mullets (Mugilidae). *bioRxiv*, 019075.

Santini, F., Carnevale, G., & Sorenson, L. (2014). First multi-locus timetree of seabreams and porgies (Percomorpha: Sparidae). *Italian Journal of Zoology*, 81(1), 55-71.

Sá-Pinto, A., Branco, M., Harris, D.J., and Alexandrino, P. (2005). Phylogeny and phylogeography of the genus Patella based on mitochondrial DNA sequence data. *Journal of Experimental Marine Biology and Ecology*, 325(1), 95-110.

Sá-Pinto, A., Branco, M., Sayanda, D., and Alexandrino, P. (2008). Patterns of colonization, evolution and gene flow in species of the genus Patella in the Macaronesian Islands. *Molecular ecology*, 17(2), 519-532.

Sato, T. (1971). A revision of the Japanese sparoid fishes of the genus Lethrinus. *Journal of the Facility of Science: University of Tokyo*, 4(12), 2.

Savage, A.E., & Zamudio, K.R. (2016). Adaptive tolerance to a pathogenic fungus drives major histocompatibility complex evolution in natural amphibian populations. *Proceedings of the Royal Society B: Biological Sciences*, 283(1827), 20153115.

Savin, S.M., Douglas, R.G., Keller, G., Killingley, J.S., Shaughnessy, L., Sommer, M.A., ... and Woodruff, F. (1981). Miocene benthic foraminiferal isotope records: a synthesis. *Marine Micropaleontology*, 6(4), 423-450.

Sbisà, E., Tanzariello, F., Reyes, A., Pesole, G., and Saccone, C. (1997). Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene*, 205(1), 125-140.

Sbrocco, E.J. (2014). Paleo-MARSPEC: gridded ocean climate layers for the mid-Holocene and Last Glacial Maximum. *Ecology*, 95, 1710-1710.

Sbrocco, E.J., and Barber, P.H. (2013). MARSPEC: ocean climate layers for marine spatial ecology. *Ecology*, 94, 979-979.

Schiebel, R., Waniek, J., Zeltner, A., and Alves, M. (2002). Impact of the Azores Front on the distribution of planktic foraminifers, shelled gastropods, and coccolithophorids. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(19), 4035-4050.

Schneider, T., Bischoff, T., and Haug, G.H. (2014). Migrations and dynamics of the intertropical convergence zone. *Nature*, 513(7516), 45.

Schouten, M.W., de Ruijter, W.P., Van Leeuwen, P.J., and Ridderinkhof, H. (2003). Eddies and variability in the Mozambique Channel. *Deep Sea Research Part II: Topical Studies in Oceanography*, 50(12), 1987-2003.

Schulter, D. (2001). Ecology and the origin of speciation. *TREE*, 16, 372-380.

Schultz, L.P. (1946). A revision of the genera of mullets, fishes of the family Mugilidae, with descriptions of three new genera. *Proceedings of the United States National Museum*.

Scott, G.R., Schulte, P.M., Egginton, S., Scott, A.L., Richards, J.G., and Milsom, W.K. (2010). Molecular evolution of cytochrome c oxidase underlies high-altitude adaptation in the bar-headed goose. *Molecular biology and evolution*, 28(1), 351-363.

Seehausen, O., Van Alphen, J.J., and Witte, F. (1997). Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science*, *277*(5333), 1808-1811.

Seehausen, O.L.E., Takimoto, G., Roy, D., and Jokela, J. (2008). Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular ecology*, 17(1), 30-44.

Seixas, F. A., Boursot, P., and Melo-Ferreira, J. (2018). The genomic impact of historical hybridization with massive mitochondrial DNA introgression. *Genome biology*, 19(1), 91.

Selkoe, K.A., Gaines, S.D., Caselle, J.E., and Warner, R.R. (2006). Current shifts and kin aggregation explain genetic patchiness in fish recruits. *Ecology*, 87(12): 3082–3094.

Semina, A.V., Polyakova, N.E., Makhotkin, M.A., and Brykov, V.A. (2007). Mitochondrial DNA divergence and phylogenetic relationships in mullets (Pisces: Mugilidae) of the Sea of Japan and the Sea of Azov revealed by PCR-RFLP-analysis. *Russian Journal of Marine Biology*, 33(3), 187-192.

Senou, H. (1988). Phylogenetic interrelationships of the mullets (Pisces: Mugilidae). *Unpublished PhD Thesis, University of Tokyo, Japan (in Japanese)*.

Seret, B., and Opic, P. (1986). Poissons de mer de l'Ouest Africain Tropical. Initiations-Documentations Techniques 49. Paris: Office de la Recherche Scientifique et Technique Outre-mer, p. 450.

Shadel, G.S., and Clayton, D.A. (1997). Mitochondrial DNA maintenance in vertebrates. *Annual review of biochemistry*, 66(1), 409-435.

Shannon, L.V. (1985). The Benguela ecosystem Part 1. Evolution of the Benguela, physical features and processes. *Oceanogr. Mar. Biol. Ann. Rev.*, *23*, 105-182.

Shannon, L.V., Lutjeharms, J.R.E., and Agenbag, J.J. (1989). Episodic input of sub-Antarctic water into the Benguela region. *South African Journal of Science*, 85(5), 317-322.

Shaw, P.W., Hendrickson, L., McKeown, N.J., Stonier, T., Naud, M.J., and Sauer, W.H.H. (2010). Discrete spawning aggregations of loliginid squid do not represent genetically distinct populations. *Marine Ecology Progress Series*, 408, 117-127.

Shaw, P.W., Pierce, G.J., and Boyle, P.R. (1999a). Subtle population structuring within a highly vagile marine invertebrate, the veined squid Loligo forbesi, demonstrated with microsatellite DNA mark-

ers. *Molecular Ecology*, 8(3), 407-417.

Shaw, P.W., Turan, C., Wright, J.M., O'connell, M., and Carvalho, G.R. (1999b). Microsatellite DNA analysis of population structure in Atlantic herring (*Clupea harengus*), with direct comparison to allozyme and mtDNA RFLP analyses. *Heredity*, 83(4): 490–499.

Shen, K.N., Jamandre, B.W., Hsu, C.C., Tzeng, W.N., and Durand, J.D. (2011). Plio-Pleistocene sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the globally-distributed flathead mullet Mugil cephalus. *BMC Evolutionary Biology*, 11(1), 83.

Sherman, K. (2006). The large marine ecosystem network approach to WSSD targets. *Ocean and coastal management*, 49(9-10), 640-648.

Shevenell, A.E., Kennett, J.P., and Lea, D.W. (2004). Middle Miocene southern ocean cooling and Antarctic cryosphere expansion. *Science*, 305(5691), 1766-1770.

Shillington, F.A., Reason, C.J.C., Rae, C.D., Florenchie, P., and Penven, P. (2006). 4 Large scale physical variability of the Benguela Current Large Marine Ecosystem (BCLME). In *Large marine ecosystems* (Vol. 14, pp. 49-70). Elsevier.

Siedler, G.,N. Zangenberg, R. Onken, and A. Morliere (1992), Seasonal changes in the tropical Atlantic Ocean: Observations and simulations of the Guinea Dome, *Journal of Geophysical Research*, 97, 703 – 715

Siesser, W.G. (1980). Late Miocene origin of the Benguela upswelling system off northern Namibia. *Science*, 208(4441), 283-285..

Siesser, W.G., and Dingle, R.V. (1981). Tertiary sea-level movements around southern Africa. *Journal of Geology*, 89, 523-536.

Silva, G., Cunha, R.L., Ramos, A., and Castilho, R. (2017). Wandering behaviour prevents inter and intra oceanic speciation in a coastal pelagic fish. *Scientific reports*, 7(1), 2893.

Silva, I.C., Mesquita, N., and Paula, J. (2010a). Genetic and morphological differentiation of the mangrove crab Perisesarma guttatum (Brachyura: Sesarmidae) along an East African latitudinal gradient. *Biological Journal of the Linnean Society,* 99, 28-46.

Silva, I.C., Mesquita, N., and Paula, J. (2010b). Lack of population structure in the fiddler crab Uca annulipes along an East African latitudinal gradient: genetic and morphometric evidence. *Marine Biology,* 157, 1113-1126.

Slatkin, M. (1993). Isolation by distance in equilibrium and nonequilibrium populations. *Evolution,* 47, 264-279.

Slatkin, M., and Hudson, R.R. (1991). Pairwise comparisons of mitochondrial-DNA sequences in stable and exponentially growing populations. *Genetics,* 129, 555-562.

Smith, J.L.B. (1959). Fishes of the family Lethrinidae from the western Indian Ocean. *Icthyological Bulletin Rhodes University,* 17, 285-295

Smith, J.L.B., Smith, M.M., and Heemstra, P.C. (2003). Smiths' sea fishes. *Struik*.

Smith, P.J., Francis, R.I.C.C., and McVeagh, M. (1991). Loss of genetic diversity due to fishing pressure. *Fisheries Research*, 10(3-4), 309-316.

Smith-Vaniz, W.F., Sidibe, A., Nunoo, F., Lindeman, K., Williams, A.B., Quartey, R., Camara, K., Carpenter, K.E., Montiero, V., de Morais, L., Djiman, R., Sylla, M. and Sagna, A. (2015). *Trachurus trachurus*. In *IUCN Red List of Threatened Species* Version 2015:e.T198647A43157137. Available at http://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T198647A43157137.en (last accessed 4 July 2019).

Snyder, M.A., Sloan, L.C., Diffenbaugh, N.S., and Bell, J.L. (2003). Future climate change and upwelling in the California Current. *Geophysical Research Letters*, 30(15).

Soekoe, M. (2016). *Adaptations in allopatric populations of Triakis megalopterus isolated by the Benguela Current. Steps towards understanding evolutionary processes affecting regional biodiversity* (Doctoral dissertation, PhD dissertation, Rhodes University, South Africa).

Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M.A.X., ... and Martin, K.D. (2007). Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *AIBS Bulletin*, 57(7), 573-583.

Stamatis, C., Triantafyllidis, A., Moutou, K.A., and Mamuris, Z. (2004). Mitochondrial DNA variation in Northeast Atlantic and Mediterranean populations of Norway lobster, Nephrops norvegicus. *Molecular Ecology*, 13(6), 1377-1390.

Stander, G.H., and De Decker, A.H.B. (1969). *Some physical and biological aspects of an oceanographic anomaly off South West Africa in 1963*., *Investl Rep. Div. Sea Fish. S. Afr*, 81, 46 pp

Stephens, M., & Donnelly, P. (2003). A comparison of bayesian methods for haplotype reconstruction from population genotype data. *The American Journal of Human Genetics, 73*(5), 1162-1169.

Stephens, M., Smith, N.J., & Donnelly, P. (2002). A new statistical method for haplotype reconstruction from population data. *The American Journal of Human Genetics*, 68(4), 978-989.

Stepien, C.A., and Rosenblatt, R.H. (1996). Genetic divergence in antitropical pelagic marine fishes (*Trachurus, Merluccius,* and *Scomber*) between North and South America. *Copeia*, 586-598.

Sternkopf, V., Liebers-Helbig, D., Ritz, M.S., Zhang, J., Helbig, A.J., and de Knijff, P. (2010). Introgressive hybridization and the evolutionary history of the herring gull complex revealed by mitochondrial and nuclear DNA. *BMC Evolutionary Biology*, 10(1), 348.

Stransky, C., Murta, A.G., Schlickeisen, J., and Zimmermann, C. (2008). Otolith shape analysis as a tool for stock separation of horse mackerel (*Trachurus trachurus*) in the Northeast Atlantic and Mediterranean. *Fisheries Research*, 89(2), 159-166.

Straughan, D.J., and Lehman, N. (2000). Genetic differentiation among Oregon lake populations of the Daphnia pulex species complex. *Journal of Heredity*, 91(1), 8-17.

Summerhayes, C.P., Kroon, D., Rosell-Melé, A., Jordan, R.W., Schrader, H.J., Hearn, R., ... and Eglinton, G. (1995). Variability in the Benguela Current upwelling system over the past 70,000 years. *Progress in oceanography*, 35(3), 207-251.

Svendsen, J.I., Alexanderson, H., Astakhov, V.I., Demidov, I., Dowdeswell, J.A., Funder, S., ... and Hubberten, H.W. (2004). Late Quaternary ice sheet history of northern Eurasia. *Quaternary Science Reviews*, 23 (11-13), 1229-1271.

Swart, N.C., Lutjeharms, J.R.E., Ridderinkhof, H., and De Ruijter, W.P.M. (2010). Observed characteristics of Mozambique Channel eddies. *Journal of Geophysical Research: Oceans*, 115(C9).

Switshenskaja, A.A., (1973). Iskopemye kefalobraznye SSSR. Trudy Paleontologicheskogo Instituta Akademii Nauk SSSR 138: 1–64 [in Russian].

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), 585-595.

Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*, 10(3), 512-526.

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A., and Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proceedings of the National Academy of Sciences*, 109 (47), 19333-19338.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular biology and evolution*, 24(8), 1596-1599.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution*, 30(12), 2725-2729.

Tamura, K., Tao, Q., and Kumar, S. (2018). Theoretical foundation of the RelTime method for estimating divergence times from variable evolutionary rates. *Molecular biology and evolution*, 35(7), 1770-1782.

Tang, Q., Liu, H., Mayden, R., and Xiong, B. (2006). Comparison of evolutionary rates in the mitochondrial DNA cytochrome b gene and control region and their implications for phylogeny of the Cobitoidea (Teleostei: Cypriniformes). *Molecular phylogenetics and evolution*, 39(2), 347-357.

Taunton-Clark, J., and Shannon, L.V. (1988). Annual and interannual variability in the South-East Atlantic during the 20th century. *South African Journal of Marine Science*, 6(1), 97-106.

Taylor, A.L., McKeown, N.J., and Shaw, P.W. (2012). Molecular identification of three co-occurring and easily misidentified octopus species using PCR-RFLP techniques. *Conservation Genetics Resources*, 4(4): 885–887.

Taylor, E.B., Boughman, J.W., Groenenboom, M., Sniatynski, M., Schluter, D., and Gow, J.L. (2006). Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (Gasterosteus aculeatus) species pair. *Molecular Ecology*, 15(2), 343-355.

Taylor, M.F., McKechnie, S.W., Pierce, N., and Kreitman, M. (1993). The lepidopteran mitochondrial control region: structure and evolution. *Molecular Biology and Evolution*, 10(6), 1259-1272.

Templeton, A.R. (1981). Mechanisms of speciation-a population genetic approach. *Annual review of Ecology*

and Systematics, 12(1), 23-48.

Templeton, A.R. (1996). Experimental evidence for the genetic-transilience model of speciation. *Evolution*, 50 (2), 909-915.

Templeton, A.R. (2008). The reality and importance of founder speciation in evolution. *Bioessays*, 30(5), 470-479.

Teske, P.R., Cherry, M.I., and Matthee, C.A. (2004). The evolutionary history of seahorses (Syngnathidae: Hippocampus): molecular data suggest a West Pacific origin and two invasions of the Atlantic Ocean. *Molecular Phylogenetics and Evolution*, 30(2), 273-286.

Teske, P.R., Froneman, P.W., Barker, N.P., and McQuaid, C.D. (2007a). Phylogeographic structure of the caridean shrimp Palaemon peringueyi in South Africa: further evidence for intraspecific genetic units associated with marine biogeographic provinces. *African Journal of Marine Science*, 29(2), 253-258.

Teske, P.R., Oosthuizen, A., Papadopoulos, I., and Barker, N.P. (2007b). Phylogeographic structure of Octopus vulgaris in South Africa revisited: identification of a second lineage near Durban harbour. *Marine Biology*,151(6), 2119-2122.

Teske, P.R., Papadopoulos, I., Newman, B.K., Dworschak, P.C., McQuaid, C.D., and Barker, N.P. (2008). Oceanic dispersal barriers, adaptation and larval retention: an interdisciplinary assessment of potential factors maintaining a phylogeographic break between sister lineages of an African prawn. *BMC Evolutionary Biology*, 8(1), 341.

Teske, P.R., Von der Heyden, S., McQuaid, C.D., and Barker, N.P. (2011). A review of marine phylogeography in southern Africa. *South African Journal of Science*, 107(5-6), 43-53.

Therkildsen, N.O., Hemmer-Hansen, J., Hedeholm, R.B., Wisz, M.S., Pampoulie, C., Meldrup, D., ... and Nielsen, E.E. (2013). Spatiotemporal SNP analysis reveals pronounced biocomplexity at the northern range margin of A tlantic cod G adus morhua. *Evolutionary Applications*, 6(4), 690-705.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, 22(22), 4673-4680.

Thomson, J.M. (1966). The grey mullets. *Oceanogr Mar Biol Ann Rev*, 4, 301-335.

Thomson, J.M. (1997). The Mugilidae of the world. *Mem Qld Mus*, 43, 457–562

Thorpe, J.P., Solé-Cava, A.M., and Watts, P.C. (2000). Exploited marine invertebrates: genetics and fisheries. In *Marine genetics* (pp. 165-184). Springer, Dordrecht.

Tieleman, B.I., Versteegh, M.A., Fries, A., Helm, B., Dingemanse, N.J., Gibbs, H.L., and Williams, J.B. (2009). Genetic modulation of energy metabolism in birds through mitochondrial function. *Proceedings of the Royal Society B: Biological Sciences*, 276(1662), 1685-1693.

Tigano, A., and Friesen, V.L. (2016). Genomics of local adaptation with gene flow. *Molecular ecology*, 25(10), 2144-2164.

Tillett, B.J., Field, I.C., Bradshaw, C.J., Johnson, G., Buckworth, R.C., Meekan, M.G., and Ovenden, J. (2012). Accuracy of species identification by fisheries observers in a north Australian shark fishery. *Fisheries Research*, 127: 109–115.

Tintore, J., La Violette, P.E., Blade, I., and Cruzado, A. (1988). A study of an intense density front in the eastern Alboran Sea: the Almeria–Oran front. *Journal of Physical Oceanography*, 18(10), 1384-1397.

Tolley, K.A., Groeneveld, J.C., Gopal, K., and Matthee, C.A. (2005). Mitochondrial DNA panmixia in spiny lobster Palinurus gilchristi suggests a population expansion. *Marine Ecology Progress Series,* 297, 225-231.

Tomita, K., Yokobori, S.I., Oshima, T., Ueda, T., and Watanabe, K. (2002). The cephalopod Loligo bleekeri mitochondrial genome: multiplied noncoding regions and transposition of tRNA genes. *Journal of Molecular Evolution*, 54(4), 486-500.

Trape, S., and Durand, J.D. (2011). First record of Mugil capurrii (Mugilidae, Perciformes) in the Gulf of Guinea. *Journal of fish biology*, 78(3), 937-940.

Trape, S., Harrison, I.J., Diouf, P.S., and Durand, J.D. (2012). Redescription of Liza bandialensis (Teleostei: Mugilidae) with an identification key to mullet species of Eastern Central Atlantic. *Comptes rendus biologies*, 335(2), 120-128.

Triantafyllidis, A., Apostolidis, A.P., Katsares, V., Kelly, E., Mercer, J., Hughes, M., ... and Triantaphyllidis, C. (2005). Mitochondrial DNA variation in the European lobster (Homarus gammarus) throughout the

range. *Marine Biology*, 146(2), 223-235.

Tringali, M.D., & Wilson, R.R. (1993). Differences in haplotype frequencies of mt-DNA of the Spanish sardine Sardinella aurita between specimens from the eastern Gulf of Mexico and southern Brazil. *Fish Bull*, 91, 362-370

Turan, C., and Yaglioglu, D. (2010). Population identification of common cuttlefish (Sepia officinalis) inferred from genetic, morphometric and cuttlebone chemistry data in the NE Mediterranean Sea. *Scientia Marina*, 74(1), 77-86.

Turan, C., Caliskan, M., and Kucuktas, H. (2005). Phylogenetic relationships of nine mullet species (Mugilidae) in the Mediterranean Sea. *Hydrobiologia*, 532(1-3), 45.

Turelli, M., Barton, N.H., and Coyne, J.A. (2001). Theory and speciation. *Trends in ecology and evolution*, 16 (7), 330-343.

Turner, T.F., Dowling, T.E., Marsh, P.C., Kesner, B.R., and Kelsen, A.T. (2007). Effective size, census size, and genetic monitoring of the endangered razorback sucker, *Xyrauchen texanus*. *Conservation Genetics*, 8 (2): 417–425.

Turpie, J.K., Beckley, L.E., and Katua, S.M. (2000). Biogeography and the selection of priority areas for conservation of South African coastal fishes.*Biological Conservation*, 92(1), 59-72.

Vai, G.B. (1997). Cyclostratigraphic estimate of the Messinian stage duration. In *Miocene stratigraphy: an integrated approach* (Vol. 15, pp. 463-476). Elsevier Amsterdam.

Valentine, J.W. (1984). Neogene marine climate trends: implications for biogeography and evolution of the shallow-sea biota. *Geology*, 12(11), 647-650.

Van Camp, L., Nykjaer, L., Mittelstaedt, E., and Schlittenhardt, P. (1991). Upwelling and boundary circulation off Northwest Africa as depicted by infrared and visible satellite observations. *Progress in Oceanography*, 26(4), 357-402.

Van der Elst, R., Everett, B., Jiddawi, N., Mwatha, G., Afonso, P.S., and Boulle, D. (2005). Fish, fishers and fisheries of the Western Indian Ocean: their diversity and status. A preliminary assessment. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 363(1826), 263-284.

Van der Elst, R.P., Groeneveld, J.C., Baloi, A.P., Marsac, F., Katonda, K.I., Ruwa, R.K., and Lane, W.L. (2009). Nine nations, one ocean: A benchmark appraisal of the South Western Indian Ocean Fisheries Project (2008–2012). *Ocean and Coastal Management*, 52(5): 258–267.

Van der Kooij, J., Engelhard, G.H., and Righton, D.A. (2016). Climate change and squid range expansion in the North Sea. *Journal of Biogeography*, 43(11), 2285-2298.

Van Der Vyver, J.S.F., Sauer, W.H.H., McKeown, N.J., Yemane, D., Shaw, P.W., and Lipinski, M.R. (2016). Phenotypic divergence despite high gene flow in chokka squid Loligo reynaudii (Cephalopoda: Loliginidae): implications for fishery management. *Journal of the Marine Biological Association of the United Kingdom*, 96(7), 1507-1525.

Van Herwerden, L., Benzie, J., and Davies, C. (2003). Microsatellite variation and population genetic structure of the red throat emperor on the Great Barrier Reef. *Journal of Fish Biology,* 62, 987-999.

Van Herwerden, L., Benzie, J., Peplow, L., and Davies, C. (2000b). Microsatellite markers for coral trout (*Plectropomus laevis*) and red throat emperor (*Lethrinus miniatus*) and their utility in other species of reef fish. *Molecular Ecology*, 9: 1929–1931.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., and Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4: 535–538.

Veiga, P., Ribeiro, J., Gonçalves, J.M.S., and Erzini, K. (2010). Quantifying recreational shore angling catch and harvest in southern Portugal (north-east Atlantic Ocean): implications for conservation and integrated fisheries management. *Journal of fish biology*, 76(9), 2216-2237.

Velema, G.J., Rosenfeld, J.S., and Taylor, E.B. (2012). Effects of invasive American signal crayfish (Pacifastacus leniusculus) on the reproductive behaviour of threespine stickleback (Gasterosteus aculeatus) sympatric species pairs. *Canadian journal of zoology*, 90(11), 1328-1338.

Viret, A., Tsaparis, D., Tsigenopoulos, C.S., Berrebi, P., Sabatini, A., Arculeo, M., ... & Caill-Milly, N. (2018). Absence of spatial genetic structure in common dentex (Dentex dentex Linnaeus, 1758) in the Mediterranean Sea as evidenced by nuclear and mitochondrial molecular markers. *PLoS one*, 13(9), e0203866.

Visram, S., Yang, M.C., Pillay, R.M., Said, S., Henriksson, O., Grahn, M., and Chen, C.A. (2010). Genetic connectivity and historical demography of the blue barred parrotfish (Scarus ghobban) in the western Indian Ocean. *Marine Biology,* 157, 1475-1487.

Von der Heyden, S. (2009). Why do we need to integrate population genetics into South African marine protected area planning?. *African Journal of Marine Science*, 31(2), 263-269.

von der Heyden, S., Prochazka, K.I.M., and Bowie, R.C. (2008). Significant population structure and asymmetric gene flow patterns amidst expanding populations of Clinus cottoides (Perciformes, Clinidae): application of molecular data to marine conservation planning in South Africa. *Molecular ecology*, 17(22), 4812-4826.

Vonlanthen, P., Bittner, D., Hudson, A.G., Young, K.A., Müller, R., Lundsgaard-Hansen, B., ... and Seehausen, O. (2012). Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature*, 482 (7385), 357.

Wagner, T. (2002). Late Cretaceous to early Quaternary organic sedimentation in the eastern Equatorial Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 179(1), 113-147.

Waliser, D.E., and Gautier, C. (1993). A satellite-derived climatology of the ITCZ. *Journal of Climate*, 6(11), 2162-2174.

Waliser, D.E., and Jiang, X. (2014). Tropical meteorology: Intertropical convergence zone. *Encycl. Atmos. Sci*.

Waples, R.S., Grewe, P.M., Bravington, M.W., Hillary, R., & Feutry, P. (2018). Robust estimates of a high Ne/N ratio in a top marine predator, southern bluefin tuna. *Science advances*, 4(7), eaar7759.

Waples, R.S. (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, 89(5), 438-450.

Waples, R.S. (2002). Evaluating the effect of stage-specific survivorship on the N-e/N ratio. *Molecular Ecology,* 11, 1029-1037.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., and Hebert, P.D. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* 360(1462), 1847-1857.

Warnock, R.C., Yang, Z., and Donoghue, P.C. (2011). Exploring uncertainty in the calibration of the molecular clock. *Biology letters*, rsbl20110710.

Warnock, R.C., Yang, Z., and Donoghue, P.C. (2017). Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. *Proceedings of the Royal Society of Biology*, 284(1857), 20170227.

Weinberg, J.R., Starczak, V.R., and Jörg, D. (1992). Evidence for rapid speciation following a founder event in the laboratory. *Evolution*, 46(4), 1214-1220.

Weir, B.S., and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population-structure. *Evolution,* 38, 1358-1370.

Welch, J.J. and Bromham, L. (2005) Molecular dating when rates vary. *Trends in Ecology and Evolution*, 20, 320–327.

Wellenreuther, M., and Bernatchez, L. (2018). Eco-evolutionary genomics of chromosomal inversions. *Trends in ecology and evolution*. 33(6), 427-440.

Wellenreuther, M., Mérot, C., Berdan, E., and Bernatchez, L. (2019). Going beyond SNP s: the role of structural genomic variants in adaptive evolution and species diversification. *Molecular ecology*. 28(6), 1203-1209.

Wernberg, T., Bennett, S., Babcock, R.C., De Bettignies, T., Cure, K., Depczynski, M., ... and Harvey, E.S. (2016). Climate-driven regime shift of a temperate marine ecosystem. *Science*, 353(6295), 169-172.

White, B.N. (1986). The isthmian link, antitropicality and American biogeography: distributional history of the Atherinopsinae (Pisces: Atherinidae). *Systematic Biology*, 35(2), 176-194.

Whiteley, A.R., Spruell, P., and Allendorf, F.W. (2004). Ecological and life history characteristics predict population genetic divergence of two salmonids in the same landscape. *Molecular Ecology*, 13(12), 3675-3688.

Whitfield, A.K. (1998). Biology and ecology of fishes in southern African estuaries. *Ichthyological Monographs of the J.L.B. Smith Institute of Ichthyology* No. 2. 223 pp.

Whitfield, A.K., Panfili, J., and Durand, J.D. (2012). A global review of the cosmopolitan flathead mullet Mugil cephalus Linnaeus 1758 (Teleostei: Mugilidae), with emphasis on the biology, genetics, ecology and

fisheries aspects of this apparent species complex. *Reviews in Fish Biology and Fisheries*, 22(3), 641-681.

Whitley, G.P. (1945). New sharks and fishes from western Australia. Pt. 2. *Aust. Zoologist.*, 11, 1-42.

Whitlock, M.C., & Mccauley, D.E. (1999). Indirect measures of gene flow and migration: FST≠ 1/(4Nm+ 1). *Heredity*, 82(2), 117-125.

Willing, E.M., Dreyer, C., and Van Oosterhout, C. (2012). Estimates of genetic differentiation measured by F ST do not necessarily require large sample sizes when using many SNP markers. *PLoS One*, 7(8), e42649.

Willis, K.J., and Niklas, K.J. (2004). The role of Quaternary environmental change in plant macroevolution: the exception or the rule?. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 359(1442), 159-172.

Winkelmann, I., Campos, P.F., Strugnell, J., Cherel, Y., Smith, P.J., Kubodera, T., ... and Norman, M. (2013). Mitochondrial genome diversity and population structure of the giant squid Architeuthis: genetics sheds new light on one of the most enigmatic marine species. *Proceedings of the Royal Society of London B: Biological Sciences*, *280*(1759), 20130273.

Winnepenninckx, B. (1993). Extraction of high molecular weight DNA from molluscs. *Trends in Genetics*, 9, 407.

Wolfram, K., Mark, F.C., John, U., Lucassen, M., and Pörtner, H.O. (2006). Microsatellite DNA variation indicates low levels of genetic differentiation among cuttlefish (Sepia officinalis L.) populations in the English Channel and the Bay of Biscay. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, 1(3), 375-383.

Woodall, L.C., Koldewey, H.J., and Shaw, P.W. (2011). Historical and contemporary population genetic connectivity of the European short-snouted seahorse Hippocampus hippocampus and implications for management. *Journal of Fish Biology*, 78(6), 1738-1756.

Woodruff, F., and Savin, S.M. (1989). Miocene deepwater oceanography. *Paleoceanography*, 4(1), 87-140.

Woolley, S., Johnson, J., Smith, M.J., Crandall, K.A., and McClellan, D.A. (2003). TreeSAAP: selection on amino acid properties using phylogenetic trees. *Bioinformatics*, 19(5), 671-672.

Wooster, W.S., Bakun, A., and McLain, D.R. (1976). Seasonal upwelling cycle along the eastern boundary of the North Atlantic. *Journal of Marine Research*, 34(2), 131-141

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, 15**:** 323–354.

Wright, S. (1978). Evolution and the Genetics of Populations. IV Variability Within and Among Populations. *University of Chicago Press*.

Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*, 26(7), 873-881.

Xavier, J.C., Allcock, A.L., Cherel, Y., Lipinski, M.R., Pierce, G.J., Rodhouse, P.G., ... and Villanueva, R. (2015). Future challenges in cephalopod research. *Journal of the Marine Biological Association of the United Kingdom*, 95(5), 999-1015.

Xavier, J.C., Peck, L.S., Fretwell, P., and Turner, J. (2016). Climate change and polar range expansions: Could cuttlefish cross the Arctic?. *Marine biology*, 163(4), 78.

Xavier, J.R., van Soest, R.W., Breeuwer, J.A., Martins, A.M., and Menken, S.B. (2010). Phylogeography, genetic diversity and structure of the poecilosclerid sponge Phorbas fictitius at oceanic islands. *Contributions to Zoology*, 79(3), 119-129.

Xia, R., Durand, J.D., and Fu, C. (2016). Multilocus resolution of Mugilidae phylogeny (Teleostei: Mugiliformes): implications for the family's taxonomy. *Molecular Phylogenetics and Evolution*, 96, 161-177.

Xia, X., & Xie, Z. (2001). DAMBE: software package for data analysis in molecular biology and evolution. *Journal of heredity*, 92(4), 371-373.

Yang, Z. (1998). On the best evolutionary rate for phylogenetic analysis. *Systematic Biology*, *47*(1), 125-133.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution*, 24 (8), 1586-1591.

Yang, Z., and Nielsen, R. (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular biology and evolution*, 19(6), 908-917.

Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A.M.K. (2000). Codon-substitution models for heterogene-

ous selection pressure at amino acid sites. *Genetics*, 155(1), 431-449.

Yemane, K., Bonnefille, R., and Faure, H. (1985). Palaeoclimatic and tectonic implications of Neogene microflora from the Northwestern Ethiopian highlands. *Nature*, 318(6047), 653.

Yokobori, S.I., Fukuda, N., Nakamura, M., Aoyama, T., and Oshima, T. (2004). Long-term conservation of six duplicated structural genes in cephalopod mitochondrial genomes. *Molecular biology and evolution*, 21(11), 2034-2046.

Yokobori, S.I., Lindsay, D.J., Yoshida, M., Tsuchiya, K., Yamagishi, A., Maruyama, T., and Oshima, T. (2007). Mitochondrial genome structure and evolution in the living fossil vampire squid, Vampyroteuthis infernalis, and extant cephalopods. *Molecular phylogenetics and evolution*, 44(2), 898-910.

Young, I.A.G., Pierce, G.J., Stowasser, G., Santos, M.B., Wang, J., and Boyle, P.R. (2006). The Moray Firth directed squid fishery. *Fisheries Research*, 78: 39–43.

Zachos, J.C., Shackleton, N.J., Revenaugh, J.S., Pälike, H., and Flower, B.P. (2001). Climate response to orbital forcing across the Oligocene-Miocene boundary. *Science*, 292(5515), 274-278.

Zane, L., Ostellari, L., Maccatrozzo, L., Bargelloni, L., Cuzin-Roudy, J., Buchholz, F., and Patarnello, T. (2000). Genetic differentiation in a pelagic crustacean (Meganyctiphanes norvegica: Euphausiacea) from the North East Atlantic and the Mediterranean Sea. *Marine Biology*, 136(2), 191-199.

Zardi, G.I., McQuaid, C.D., Teske, P.R., and Barker, N.P. (2007). Unexpected genetic structure of mussel populations in South Africa: indigenous Perna perna and invasive Mytilus galloprovincialis. *Marine Ecology Progress Series*, 337, 135-144.

Zardoya, R., Castilho, R., Grande, C., Favre-Krey, L., Caetano, S., Marcato, S., ... and Patarnello, T. (2004). Differential population structuring of two closely related fish species, the mackerel (Scomber scombrus) and the chub mackerel (Scomber japonicus), in the Mediterranean Sea. *Molecular Ecology*, 13(7), 1785-1798.

Zavatarielli, M., and Mellor, G.L. (1995). A numerical study of the Mediterranean Sea circulation. *Journal of Physical Oceanography*, 25(6), 1384-1414.

Zeidberg, L.D., and Robison, B.H. (2007). Invasive range expansion by the Humboldt squid, Dosidicus gigas, in the eastern North Pacific. *Proceedings of the National Academy of Sciences*, 104(31), 12948-12950.

Zemlak, T.S., Ward, R.D., Connell, A.D., Holmes, B.H., and Hebert, P.D.N. (2009). DNA barcoding reveals overlooked marine fishes. *Molecular Ecology Resources,* 9, 237-242.

Zenk, W., Klein, B., and Schroder, M. (1991). Cape Verde frontal zone. *Deep Sea Research Part A. Oceanographic Research Papers*, 38, S505-S530.

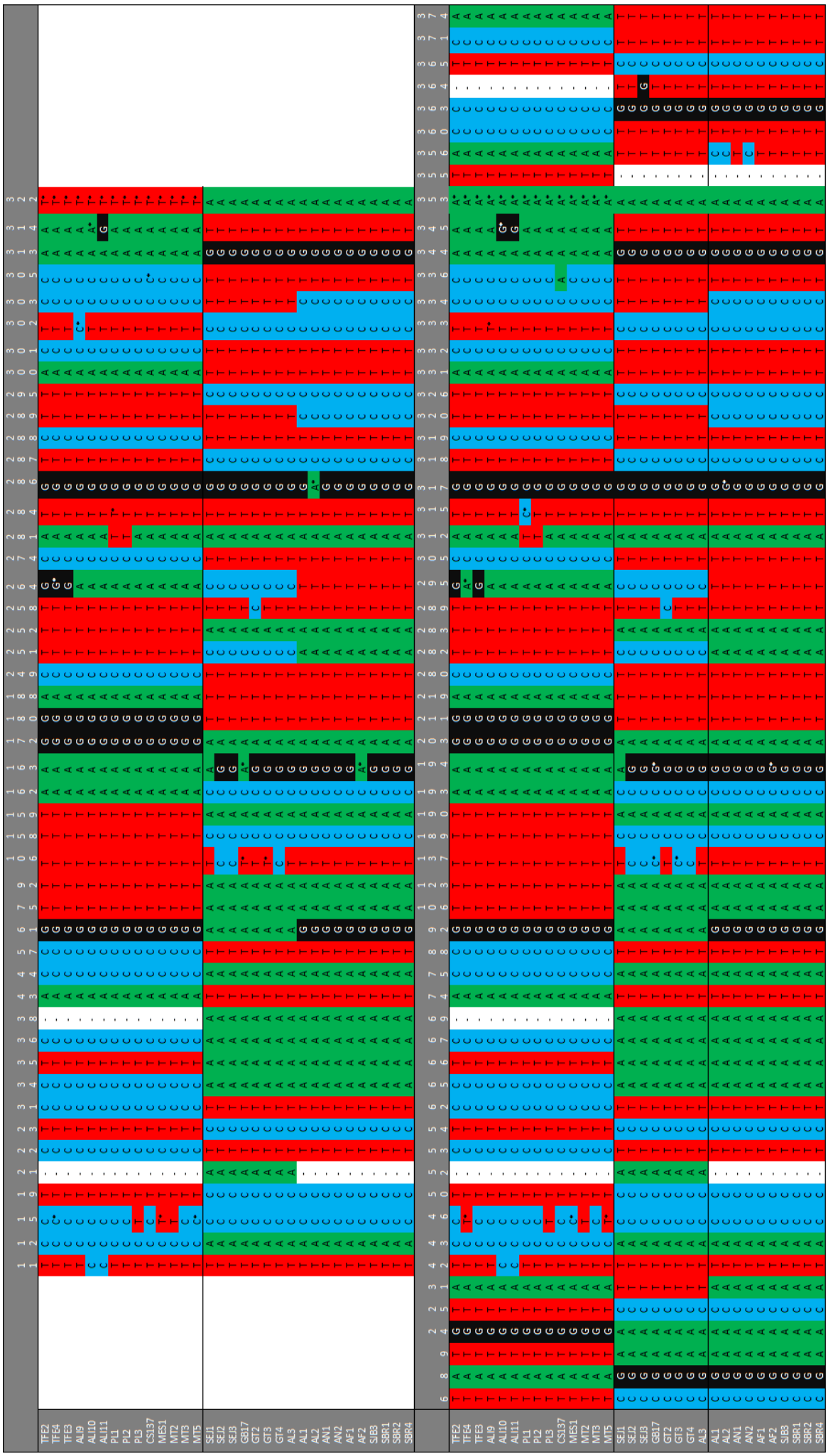
Zhang, D.X., & Hewitt, G.M. (1997). Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. *Biochemical Systematics and Ecology*, 25(2), 99-120.

Zigler, K.S., and Lessios, H.A. (2004). Speciation on the coasts of the new world: phylogeography and the evolution of bindin in the sea urchin genus Lytechinus. *Evolution*, 58(6), 1225-1241.

Zulliger, D.E., Tanner, S., Ruch, M., and Ribi, G. (2009). Genetic structure of the high dispersal Atlanto-Mediterreanean sea star Astropecten aranciacus revealed by mitochondrial DNA sequences and microsatellite loci. *Marine Biology*, 156(4), 597-610.

SUPPLEMENTARY MATERIAL FOR CHAPTER 2.1

Supplementary Table 2.1.1: Variable sites within and between NCR1 and NCR2 alignments of *S. officinalis*, *S. hierredda* and *S. vermiculata*. Variable sites at NCR1 are presented in the top table, whilst NCR2s are in the bottom table. * indicate a difference in nucleotide composition between NCRs



SUPPLEMENTARY MATERIAL FOR CHAPTER 4

Supplementary Table 4.1: *P* values from single locus by sample tests of conformance to Hardy-Weinberg equilibrium genotype proportions.

Locus	PO	MH	ME	MT	GT	AN	SAS
TmurA101	0.021	0.060	0.000	0.000	0.000	0.132	0.057
TmurA104	0.254	0.930	0.000	0.081	0.030	0.488	0.000
TmurA115	0.003	0.000	0.000	0.000	0.000	0.169	0.003
TmurB104	0.015	0.172	0.000	0.067	0.099	0.886	0.256
TmurB116	0.199	0.058	0.934	0.001	0.599	0.484	0.962
TmurC4	0.001	0.492	0.852	0.099	0.196	0.626	0.005
TT48	1.000	0.958	0.167	0.953	0.079	0.659	0.787
TT62	0.387	0.698	0.001	0.115	0.000	0.141	0.013
TT113	0.008	0.129	0.044	0.468	0.000	0.559	0.027
TT29	0.204	0.027	0.000	0.001	0.000	0.720	0.005

Supplementary Table 4.2: Estimated frequency of null alleles for each locus x sample combination

	TmurA101	TmurA104	TmurA115	TmurB104	TmurB116	TmurC4	TT48	TT62	TT113	TT29
PO	0.17	0.14	0.29	0.23	0.00	0.11	0.00	0.08	0.19	0.07
MH	0.23	0.00	0.38	0.20	0.05	0.14	0.00	0.00	0.10	0.19
ME	0.23	0.04	0.32	0.12	0.03	0.00	0.03	0.16	0.07	0.24
MT	0.19	0.05	0.33	0.15	0.04	0.07	0.02	0.14	0.05	0.13
GT	0.13	0.12	0.22	0.09	0.02	0.02	0.02	0.19	0.10	0.21
AN	0.07	0.10	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SAS	0.11	0.16	0.18	0.14	0.00	0.03	0.00	0.09	0.06	0.17

Supplementary Table 3. Pairwise *F*ST estimated from the three loci conforming to HWE i.e. no null alleles (below diagonal). Above diagonal are reported corresponding *F*ST *P* values estimated following permutation

	PO	MH	ME	MT	GT	AN	SAS
PO	-	0.312	0.469	0.362	0.282	0.016	0.005
MH	0.028	-	0.644	0.190	0.255	0.156	0.029
ME	0.013	0.012	-	0.223	0.545	0.004	0.002
MT	0.016	0.022	0.009	-	0.498	0.004	0.008
GT	0.016	0.018	0.005	0.006	-	0.003	0.005
AN	0.061	0.043	0.046	0.057	0.055	-	0.014
SAS	0.047	0.038	0.025	0.023	0.023	0.047	-

SUPPLEMENTARY MATERIAL FOR CHAPTER 5.1

Appendix

Appendix A: PCR-RFLP protocol

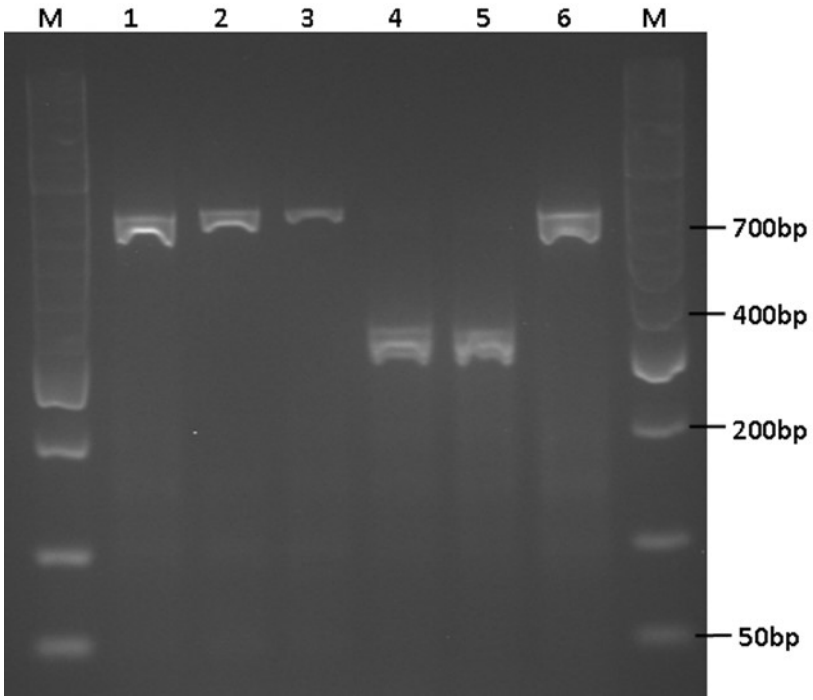


Figure A5.1.1.1: mtDNA gene digest of *Lethrinus mahsena* clades A and B and their associated fragment patterns when separated by gel electrophoresis. *M* molecular weight marker Bioline Hyperladder 50bp; *1* unrestricted PCR product *L. mahsena* clade A; *2-3* gene digest of *L. mahsena* clade A; *4-5* gene digest of *L. mahsena* clade B; *6* unrestricted PCR product *L. mahsena* clade B.

Method: In order to differentiate between the two highly divergent mtDNA COI clades observed in *L. mahsena* (see 3.1), a PCR-RFLP assay was developed. This method allowed for the high throughput screening of all *L. mahsena* individuals, thus permitting direct comparisons of mtDNA and microsatellite genotypes. Sequence alignments of the CO1 gene of both clades of *L. mahsena* were analysed using NEB cutter v2.0 (Vincze et al. 2003) and diagnostic restriction enzyme cleavage sites for the enzyme *FauI* were identified. Species specific COI primers (see methods for PCR protocol) were then used to PCR amplify a 480bp fragment of the CO1 gene. The amplified CO1 fragment was then cut in a subsequent reaction using the enzyme *FauI*. Each reaction contained 5µL of the amplified CO1 sequence, 0.1 µL of *FauI*, and 1µL of enzyme specific buffer and 3.9µL of ddH_2O. This was then incubated in a thermocycler for 2 hours at 55°C followed by 20 minutes at 65°C. Finally, individual mtDNA clades were assigned by gel electrophoresis (Fig.A1) and microsatellite genotypes assigned to mtDNA clades for subsequent analysis

SUPPLEMENTARY MATERIAL FOR CHAPTER 5.1

Supplementary Table 5.1.1: Bioclimatic variables used in species distribution modelling.

Layer Name	Layer Definition	Units	Scaling Factor	Derived from
bathymetry	Depth of the seafloor	meters	1×	SRTM30 Plus Bathymetry
biogeo01	East/West Aspect (sin(aspect in radians))	radians	100×	Bathymetry
biogeo02	North/South Aspect (cos(aspect in radians))	radians	100×	Bathymetry
biogeo03	Plan Curvature	none	10,000×	Bathymetry
biogeo04	Profile Curvature	none	10,000×	Bathymetry
biogeo05	Distance to Shore	kilometers	1×	GSHHS Coastline
biogeo06	Bathymetric Slope	degrees	10×	Bathymetry
biogeo07	Concavity	degrees	1000×	Bathymetry
biogeo08	Mean Annual SSS	psu	100×	SSS monthly climatologies
biogeo09	Minimum Monthly SSS	psu	100×	SSS monthly climatologies
biogeo10	Maximum Monthly SSS	psu	100×	SSS monthly climatologies
biogeo11	Annual Range in SSS	psu	100×	SSS monthly climatologies
biogeo12	Annual Variance in SSS	psu	10,000×	SSS monthly climatologies
biogeo13	Mean Annual SST	degrees C	100×	SST monthly climatologies
biogeo14	SST of the coldest ice-free month	degrees C	100×	SST monthly climatologies
biogeo15	SST of the warmest ice-free month	degrees C	100×	SST monthly climatologies
biogeo16	Annual Range in SST	degrees C	100×	SST monthly climatologies
biogeo17	Annual Variance in SST	degrees C	10,000×	SST monthly climatologies

Supplementary Table 5.1.2: Haplotype frequencies for 486bp of *Lethrinus mahsena* COI mtDNA (Haplotypes 4-8 belong to Clade B).

Haplotype	MaM	SP	SM	TZ	MzP	MdB	MzM	Total
1	12	0	1	12	1	30	4	33
2	0	0	1	0	0	0	0	1
3	0	0	1	0	0	0	0	1
4	0	0	1	0	0	0	0	1
5	0	1	2	0	0	0	0	3
6	0	1	1	0	0	0	0	2
7	0	6	6	0	0	0	0	12
8	0	1	0	0	0	0	0	1
9	0	0	0	1	0	0	0	1
10	0	0	0	1	1	0	0	2
11	3	0	0	2	1	0	1	7
12	2	0	0	2	0	0	1	5
13	1	0	0	0	0	0	0	1
14	0	0	0	0	0	2	0	2
Total	18	9	13	18	3	5	6	72

Supplementary Table 5.1.3: Haplotype frequencies for 510 bp of *Lethrinus harak* COI mtDNA

Haplotype	MaM	TZ	MzP	MdB	Total
1	6	9	17	28	60
2	16	7	4	13	40
3	0	1	0	0	1
4	0	1	0	0	1
5	0	0	1	1	2
Total	22	18	22	22	104

Supplementary Table 5.1.4: Kamura-2-Parameter genetic distances between (below diagonal) and within (on diagonal) putative Lethrinid species, *L. mahsena* (Moz) and *L.* mahsena (India) correspond to GENBANK samples: KJ920117.1, EF609387.1, JF493750.1, JF493751.1 and JF493752.1. Standard errors given in brackets.

	***L.mahsena* CA**	***L.mahsena* (Moz)**	***L. mahsena* CB**	***L.mahsena* (India)**	***L. atkinsoni***	***L. harak***
***L.mahsena* CA**	0.003 (0.002)					
***L.mahsena* (Moz)**	0.002 (0.001)	0.003 (0.002)				
***L. mahsena* CB**	0.015 (0.003)	0.048 (0.010)	0.002 (0.001)			
***L. mahsena* (Ind)**	0.047 (0.010)	0.047 (0.010)	0.004 (0.002)	0.006 (0.003)		
L. atkinsoni	0.046 (0.009)	0.048 (0.010)	0.040 (0.008)	0.024 (0.006)	0.001 (0.001)	
L. harak	0.091 (0.014)	0.091 (0.015)	0.092 (0.013)	0.094 (0.015)	0.100 (0.015)	0.006 (0.003)

Supplementary Table 5.1.5: Likelihood of detecting significance (from zero) of a range of F_{ST} values based on POWSIM analyses of 486bp mtDNA (COI) for *L. harak* and *L. mahsena* and 10 and 9 microsatellite loci in *Lethrinus mahsena Lethrinus harak* respectively. Comparisons were made between average (a) and minimum (m) population sizes for the respective datasets.

	COI						**msat**					
	L. harak			***L. mahsena***			***L.harak***			***L. mahsena***		
	a v a	**a v m**	**m v m**	**a v a**	**a v m**	**m v m**	**a v a**	**a v m**	**m v m**	**a v a**	**a v m**	**m v m**
F_{ST}=0	0.038	0.024	0.035	0.04	0.026	0.013	0.049	0.035	0.043	0.047	0.028	0.008
F_{ST}=0.05	0.222	0.163	0.106	0.074	0.125	0.009	1	1	1	0.94	0.659	0.233

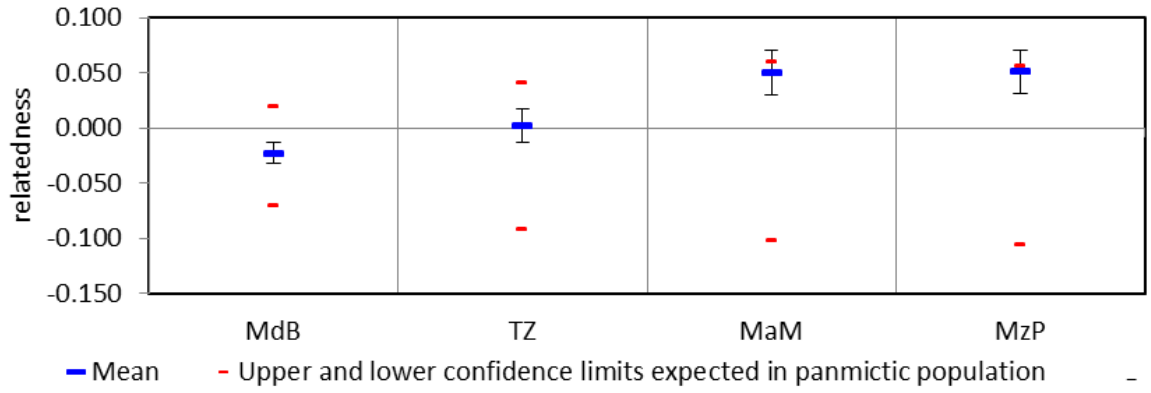
Supplementary Table 5.1.7: Microsatellite genetic diversity across 9 loci and 4 populations in *L. harak*: microsatellite sample size = N, number of alleles = N_A, Allelic richness= A_R, expected heterozygosity = H_E, observed heterozygosity = H_O and probability of deviation from Hardy Weinberg expectations = HWE.

		80	90	2.33	58	100	96	952	68	75	Total
MaM	N	39	41	42	37	43	43	43	43	43	41.556
	N_A	2	2	6	5	3	3	7	6	2	4.000
	A_R	2.000	2.000	5.524	5.000	2.858	2.982	6.858	5.840	2.000	
	H_O	0.487	0.341	0.310	0.595	0.093	0.837	0.860	0.605	0.512	0.516
	H_E	0.369	0.283	0.383	0.716	0.090	0.505	0.781	0.494	0.381	0.445
	F_{IS}	-0.310	-0.194	0.203	0.183	-0.028	-0.651	-0.091	-0.212	-0.333	-0.148
	HWE	0.080	0.573	**0.043**	**0.001**	1.000	**0.000**	0.097	0.631	**0.040**	**0.000**
TZ	N	60	60	60	57	57	55	58	57	59	58.111
	N_A	2	2	4	5	4	3	7	8	2	4.111
	A_R	2.000	1.997	3.327	4.999	4.000	3.000	6.998	7.703	2.000	
	H_O	0.333	0.100	0.050	0.544	0.596	0.545	0.690	0.667	0.288	0.424
	H_E	0.278	0.095	0.081	0.681	0.525	0.444	0.829	0.620	0.247	0.422
	F_{IS}	-0.192	-0.044	0.390	0.210	-0.128	-0.219	0.177	-0.066	-0.160	0.005
	HWE	0.340	1.000	**0.001**	**0.001**	0.698	0.167	**0.013**	0.465	0.588	**0.000**
MzP	N	44	44	43	43	44	42	43	43	44	43.333
	N_A	3	2	2	4	4	3	7	8	2	3.889
	A_R	3.000	1.997	1.860	3.982	3.682	3.000	6.860	7.842	2.000	
	H_O	0.205	0.068	0.023	0.349	0.455	0.738	0.744	0.651	0.455	0.410
	H_E	0.362	0.066	0.023	0.530	0.361	0.602	0.771	0.682	0.351	0.416
	F_{IS}	0.444	-0.024	-	0.353	-0.249	-0.215	0.047	0.057	-0.284	0.028
	HWE	**0.000**	1.000	-	**0.003**	0.485	**0.036**	0.423	0.109	0.086	**0.000**
MDB	N	97	99	99	97	100	94	96	94	96	96.889
	N_A	3	3	4	6	5	4	8	8	2	4.778
	A_R	2.381	2.929	3.201	5.147	3.726	3.394	7.811	7.318	2.000	
	H_O	0.340	0.141	0.091	0.588	0.280	0.564	0.750	0.755	0.229	0.415
	H_E	0.284	0.134	0.125	0.643	0.251	0.560	0.822	0.717	0.249	0.421
	F_{IS}	-0.193	-0.051	0.277	0.091	-0.110	-0.001	0.093	-0.049	0.085	0.017
	HWE	0.130	1.000	**0.000**	**0.000**	0.887	0.991	0.104	0.067	0.412	**0.000**
Total	N	60.000	61.000	61.000	58.500	61.000	58.500	60.000	59.250	60.500	59.972
	N_A	2.500	2.250	4.000	5.000	4.000	3.250	7.250	7.500	2.000	4.194
	A_R	2.970	2.680	4.141	5.096	4.125	3.158	7.738	7.764	2.000	
	H_O	0.341	0.163	0.118	0.519	0.356	0.671	0.761	0.669	0.371	0.441
	H_E	0.323	0.144	0.153	0.642	0.307	0.528	0.801	0.628	0.307	0.426
	F_{IS}	-0.057	-0.127	0.225	0.193	-0.161	-0.271	0.050	-0.066	-0.208	-0.047
	HWE	**0.000**	0.997	**0.000**	**0.000**	0.966	**0.000**	**0.012**	0.138	0.078	**0.000**

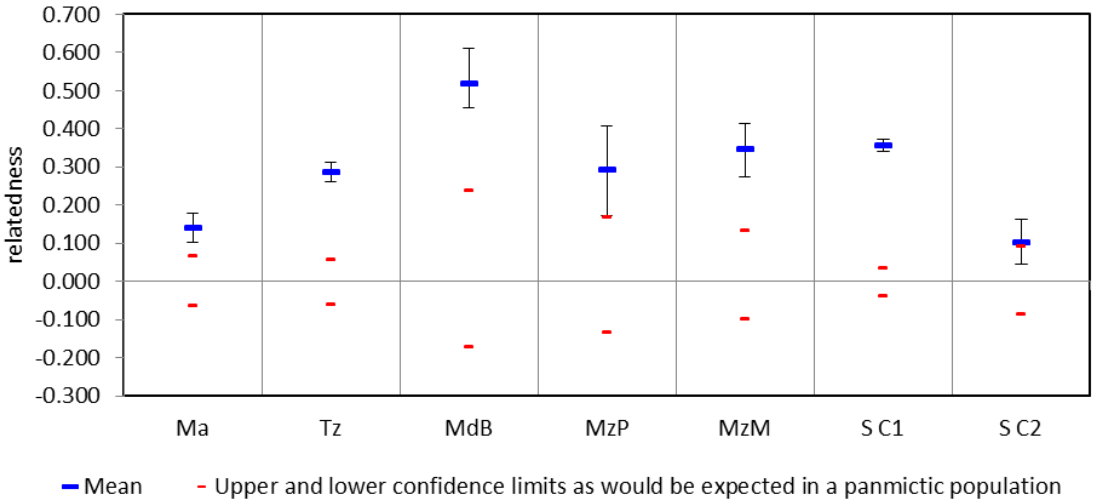
Supplementary Table 5.1.6: Microsatellite genetic diversity across 10 loci and 6 populations in *L. mahsena*: microsatellite sample size = N, number of alleles = N_A, Allelic richness= A_R, expected heterozygosity = H_E, observed heterozygosity = H_O and probability of deviation from Hardy Weinberg expectations = HWE.

		80	951	90	2.33	58	100	96	952	68	75	Total
MaM	N	19	19	16	19	17	19	18	19	19	19	18.4
	N_A	4	3	4	6	7	2	3	6	5	4	4.4
	A_R	3.200	2.596	3.113	3.573	3.889	1.995	2.392	3.386	3.023	2.137	-
	H_O	0.684	0.684	0.500	0.526	0.824	0.684	0.389	0.684	0.579	0.368	0.592
	H_E	0.652	0.506	0.662	0.693	0.689	0.494	0.406	0.591	0.568	0.316	0.558
	F_{IS}	-0.022	-0.330	0.275	0.265	-0.167	-0.361	0.070	-0.130	0.008	-0.140	-0.034
	HWE	**0.000**	0.269	0.183	0.076	0.945	0.176	0.471	0.568	0.766	1.000	**0.023**
SM	N	10	9	9	10	9	10	10	10	9	10	9.6
	N_A	3	2	4	4	2	2	4	4	7	4	3.6
	A_R	2.935	1.999	2.333	2.200	1.996	1.999	3.029	3.294	4.812	2.200	-
	H_O	0.800	0.889	0.333	0.200	0.556	0.900	0.500	0.500	0.778	0.300	0.576
	H_E	0.655	0.494	0.296	0.270	0.475	0.495	0.595	0.655	0.790	0.270	0.500
	F_{IS}	-0.171	-0.778	-0.067	0.308	-0.111	-0.800	0.211	0.286	0.074	-0.059	-0.098
	HWE	**0.002**	0.056	1.000	0.158	1.000	**0.046**	0.489	0.633	0.600	1.000	**0.044**
TZ	N	19	21	21	20	21	21	21	21	21	21	20.7
	N_A	2	3	5	5	7	3	3	6	5	5	4.4
	A_R	1.955	2.154	3.904	3.899	4.804	2.299	1.934	2.715	2.706	2.797	-
	H_O	0.474	0.667	0.714	0.500	0.714	0.476	0.190	0.429	0.571	0.619	0.535
	H_E	0.411	0.459	0.757	0.738	0.817	0.472	0.254	0.407	0.481	0.488	0.528
	F_{IS}	-0.125	-0.432	0.081	0.345	0.150	0.015	0.273	-0.029	-0.165	-0.247	0.011
	HWE	1.000	0.077	0.853	**0.018**	0.271	1.000	0.173	0.769	0.808	0.891	0.412
MzP	N	7	7	7	6	7	7	7	7	7	7	6.9
	N_A	2	3	2	2	4	3	4	5	4	2	3.1
	A_R	1.985	2.407	1.835	1.909	2.978	2.820	3.501	4.294	3.392	1.571	-
	H_O	0.571	0.429	0.000	0.000	0.429	0.857	0.857	0.857	0.429	0.143	0.457
	H_E	0.408	0.357	0.245	0.278	0.459	0.571	0.684	0.755	0.643	0.133	0.453
	F_{IS}	-0.333	-0.125	1.000	1.000	0.143	-0.440	-0.180	-0.059	0.400	-	0.070
	HWE	1.000	1.000	0.077	0.091	0.441	0.478	0.712	0.471	**0.040**	-	0.246
MzM	N	8	10	10	8	10	10	10	10	10	9	9.5
	N_A	3	3	3	4	8	3	3	5	3	1	3.6
	A_R	2.754	2.349	2.305	2.987	5.412	2.602	2.602	3.954	2.305	1.000	-
	H_O	0.250	0.400	0.400	0.250	1.000	0.600	0.700	0.700	0.200	0.000	0.450
	H_E	0.570	0.445	0.340	0.578	0.840	0.505	0.505	0.745	0.340	0.000	0.487
	F_{IS}	0.606	0.153	-0.125	0.611	-0.133	-0.137	-3.404	0.113	0.455	-	0.133
	HWE	**0.021**	1.000	1.000	**0.021**	0.733	1.000	0.505	**0.035**	0.108	-	0.053
MDB	N	5	5	5	4	5	5	5	5	5	5	4.9
	N_A	2	2	2	2	5	2	1	2	3	2	2.3
	A_R	2.000	2.000	2.000	2.000	4.578	2.000	1.000	1.978	2.800	1.800	-
	H_O	0.600	0.800	0.600	0.500	1.000	0.800	0.000	0.400	0.400	0.200	0.530
	H_E	0.420	0.480	0.420	0.375	0.760	0.480	0.000	0.320	0.580	0.180	0.402
	F_{IS}	-0.333	-0.600	-0.333	-0.200	-0.212	-0.600	-	-0.143	0.407	-	-0.214
	HWE	1.000	0.429	1.000	1.000	0.848	0.429	-	1.000	0.365	-	0.991
Total	N	11.333	11.833	11.333	11.167	11.500	12.000	11.833	12.000	11.833	11.833	11.667
	N_A	2.667	2.667	3.333	3.833	5.500	2.500	3.000	4.667	4.500	3.000	3.567
	A_R	3.146	2.896	3.173	4.045	4.589	2.286	2.860	3.674	3.634	2.189	-
	H_O	0.563	0.645	0.425	0.329	0.754	0.720	0.439	0.595	0.493	0.272	0.523
	H_E	0.520	0.457	0.453	0.488	0.673	0.503	0.407	0.579	0.567	0.231	0.488
	F_{IS}	-0.084	-0.411	0.064	0.326	-0.119	-0.431	-0.079	-0.028	0.131	-0.176	-0.081

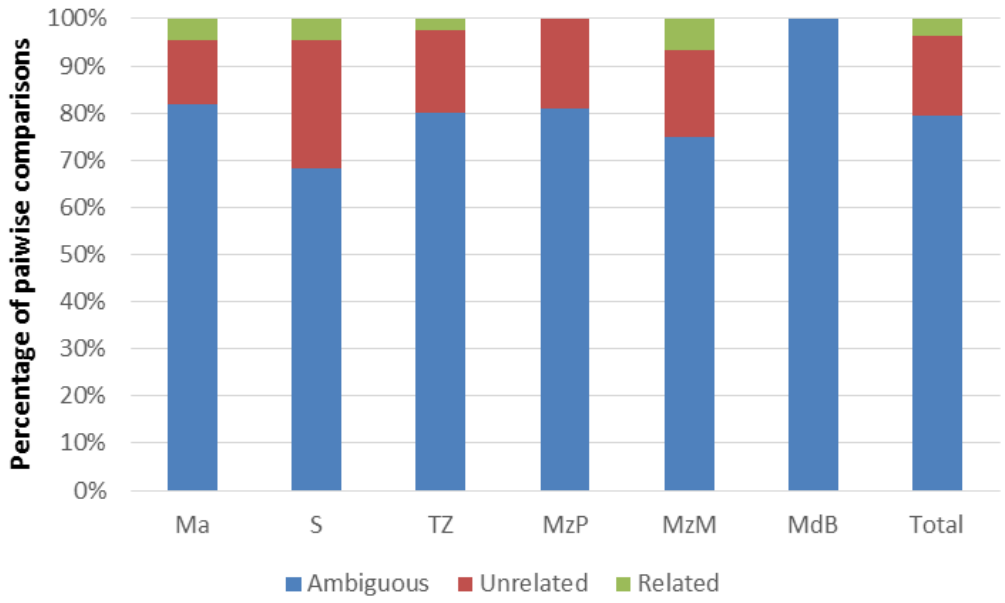
Supplementary Figure 5.1.1: Assessment of number of populations of A) *Lethrinus mahsena* and B) *Lethrinus harak* across the SWIO, estimated through Bayesian admixture implemented in BAPS.



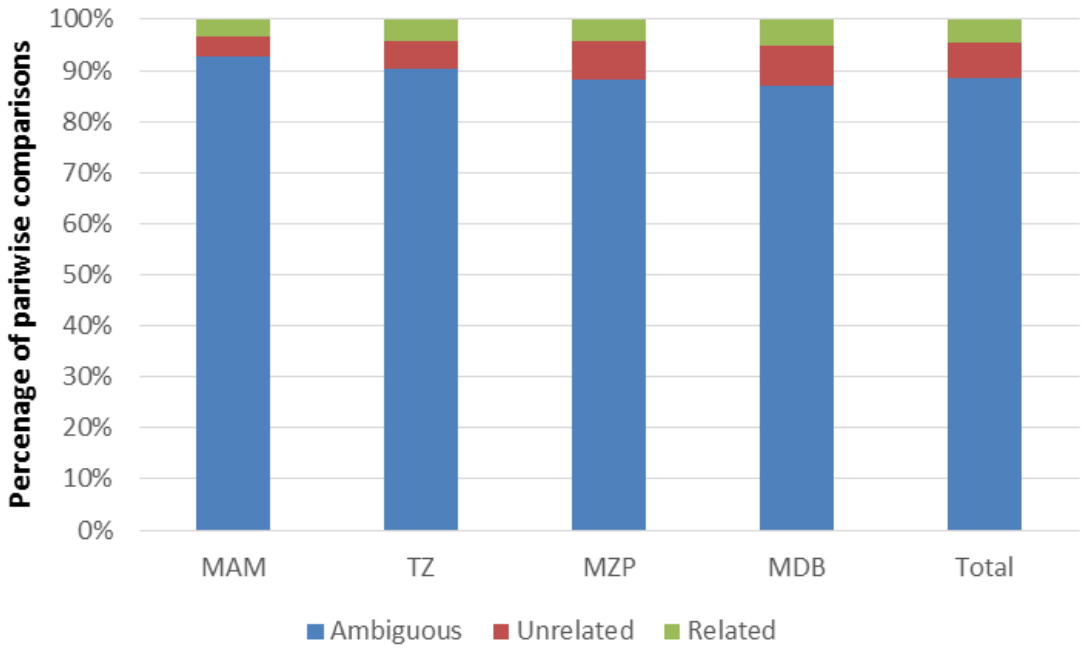
Supplementary Figure 5.1.2: Relatedness (r_{qg}) for *Lethrinus harak* samples across the SWIO.



Supplementary Figure 5.1.3: Relatedness (r_{qg}) for *Lethrinus mahsena* samples across the SWIO, Seychelles islands are pooled and separated by mtDNA clade A (S C1) and clade B (S C2).



Supplementary Figure 5.1.4: Percentage of relationships between pairs of individuals of *L. mahsena* that fell within 3 relatedness classifications: (a). unrelated (classification as U only), (b) related (classification as any combination of Half Sibling, Full Sibling, Parent-Offspring but not Unrelated), (c) ambiguous (classification as Unrelated as well as some related state). Based on 10 microsatellite loci.



Supplementary Figure 5.1.5: Percentage of relationships between pairs of individuals of *L. harak* that fell within 3 relatedness classifications: (a). unrelated (classification as U only), (b) related (classification as any combination of Half Sibling, Full Sibling, Parent-Offspring but not Unrelated), (c) ambiguous (classification as Unrelated as well as some related state). Based on 9 microsatellite loci.

SUPPLEMENTARY MATERIAL FOR CHAPTER 5.2

Supplementary Table 5.2.1: Global F_{ST} values for each of 14 polymorphic microsatellite loci and over all loci across eight populations of *Lethrinus nebulosus*, with or without null allele frequencies estimated using the ENA method.

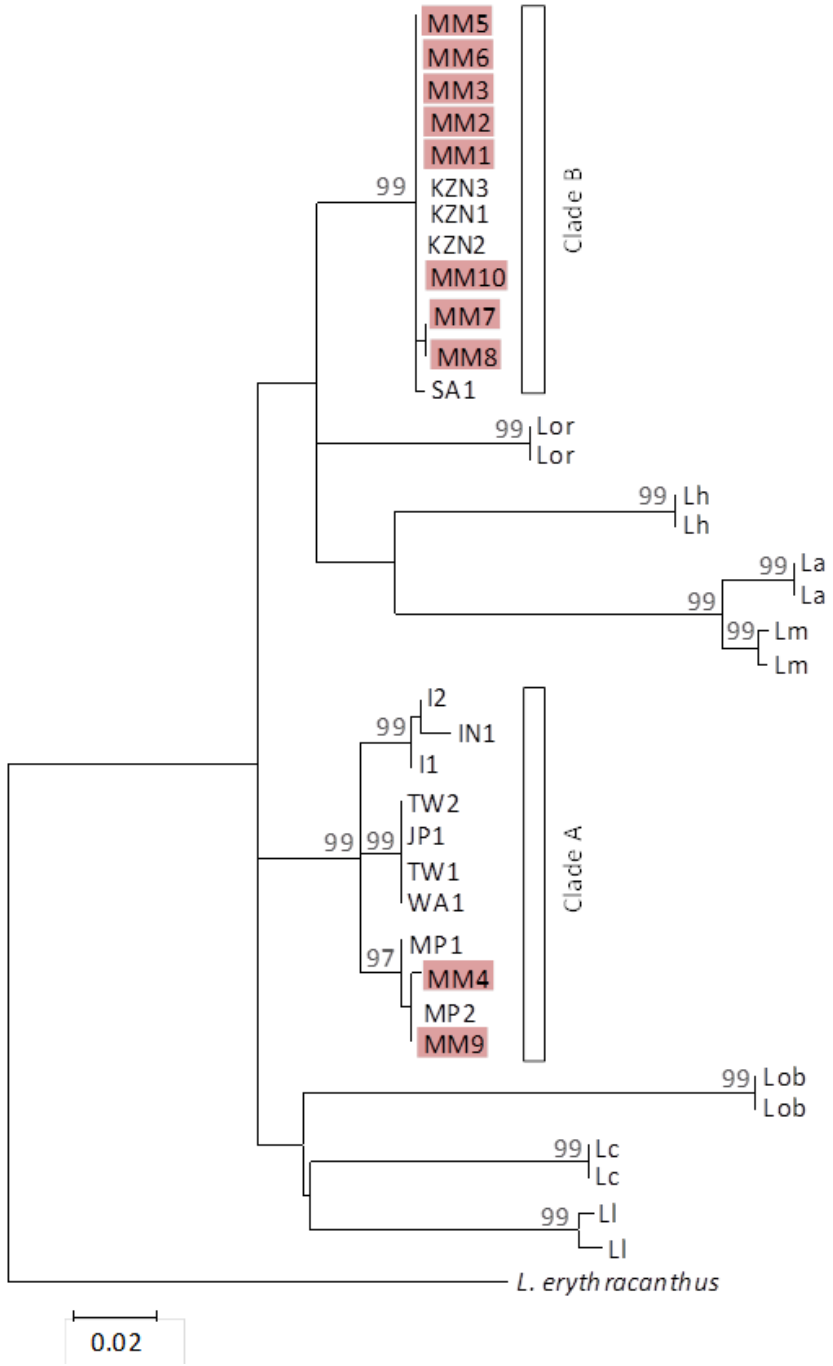
Locus	F_{ST} without ENA	F_{ST} with ENA
19RTE	0.017	0.016
23RTE	0.035	0.035
37RTE	0.069	0.070
41RTE	0.013	0.022
44RTE	0.039	0.040
58RTE	0.057	0.058
68RTE	0.022	0.020
75RTE	0.032	0.031
83RTE	0.026	0.026
90RTE	0.130	0.138
95ACRTE	0.024	0.024
96RTE	0.119	0.121
100RTE	0.025	0.023
Bst 2.33	0.074	0.077
Overall	0.046 (0.032 – 0.063)	0.047 (0.033 – 0.065)

Supplementary Table 5.2.2: Sampling regions, individual sampling localities and sample sizes (*N*) for *Lethrinus nebulosus* sampled across the SWIO.

Region	Code	Locality	N (msat)	Total (msat)	N (CR)	Total (CR)
Seychelles	S	Mahe	7	37	19	49
		Farquhar	30		30	
Kenya	K		29	29	29	29
Tanzania	TZ		30	30	41	41
North Mozambique Channel	NMC	Comoros	5	11	5	10
		Morondava, Madagascar	1			
		Nosy Bé, Madagascar	5		5	
West Madagascar	WM	Andavadoaka			14	22
		Belo sur Mer			8	
Mozambique	MZ	Zavora	4	32	4	32
		Quissico	28		28	
South Africa	SA	Bhanga Nek	3	60	3	60
		Richard's Bay	57		57	
East Madagascar	EM	St. Marie	9	12	9	13
		Toamasina	3		4	
Mauritius	MA		30	30	36	36

Supplementary Table 5.2.3: Percentage sequence divergence and p-distance between clades A , B and *Lethrinus ornatus* using 421bp of mtDNA CR and 443bp mtDNA COI.

	Sequence divergence (%)	p distance
CR		
Clade A - Clade B	11.96	0.143
COI		
Clade A - Clade B	5.64	0.057
Clade A - *L. ornatus*	5.66	0.066
Clade B - *L. ornatus*	3.16	0.056



Supplementary Figure 5.2.1: Maximum likelihood topology depicting the relationship amongst a sub-sample of Clade A and Clade B *L. nebulosus* samples from Mozambique Maputo (MM, highlighted in red) and other Lethrinid species, based on 443bp mtDNA COI. *L. nebulosus*: KZN=Kwazulu-Natal (DQ885022.1, DQ885021.1, DQ885020.1), SA=South Africa (JF493753.1), I=Iran (HQ149871.1, HQ149872.1), IN= India (KM079310.1), TW=Taiwan (KU944039.1, KU944037.1), JP= Japan (JF952783.1), WA= Western Australia (DQ885100.1), MP=Mozambique, Pomene (JF493754.1, HQ561492.1). Lor= *L.* ornatus (KF809405.1, KC970480.1), Lh= *L. harak* (JQ350084.1, HQ561476), La= *L. atkinsoni* (KP194639.1, KP194152.1), Lm=*L. mahsena* (KJ920117.1, EF609387.1), Lob= *L. obsoletus* (KP194670.1, KP194475.1), Lc= *L. crocineus* (KM079301.1, KJ920118.1), Ll= *L. lentjan* (KF930047.1, FJ237800.1). Bootstrap values are given on the branch. Branch lengths are proportional to the number of nucleotide substitutions and *L. erythracanthus* is included as an outgroup species.

SUPPLEMENTARY MATERIAL FOR CHAPTER 7
Demographic history and divergence of the *Sepia officinalis* species complex

Methods, demographic tests and timescales of divergence

823bp of concatenated mtDNA COI and cytb sequences, employed in chapter 2.3, were used to explore the timing of diversification and demographic history of the *S. officinalis* species complex. Divergence times for the major clades of the Sepia phylogeny were estimated using the RelTime (Tamura *et al.* 2012; 2013) method, implemented in MEGA v7 (see chapter 3.1.2.2 for a more comprehensive discussion of this technique). RelTime was run using the HKY+G+I substitution model and 2 fossil calibrations were applied to internal nodes based on Strugnall *et al.* (2006): A) upper bound=37MYA, lower bound=19MYA; B) upper bound=6MYA, lower bound=3MYA (see Supplementary Figure 6.1).

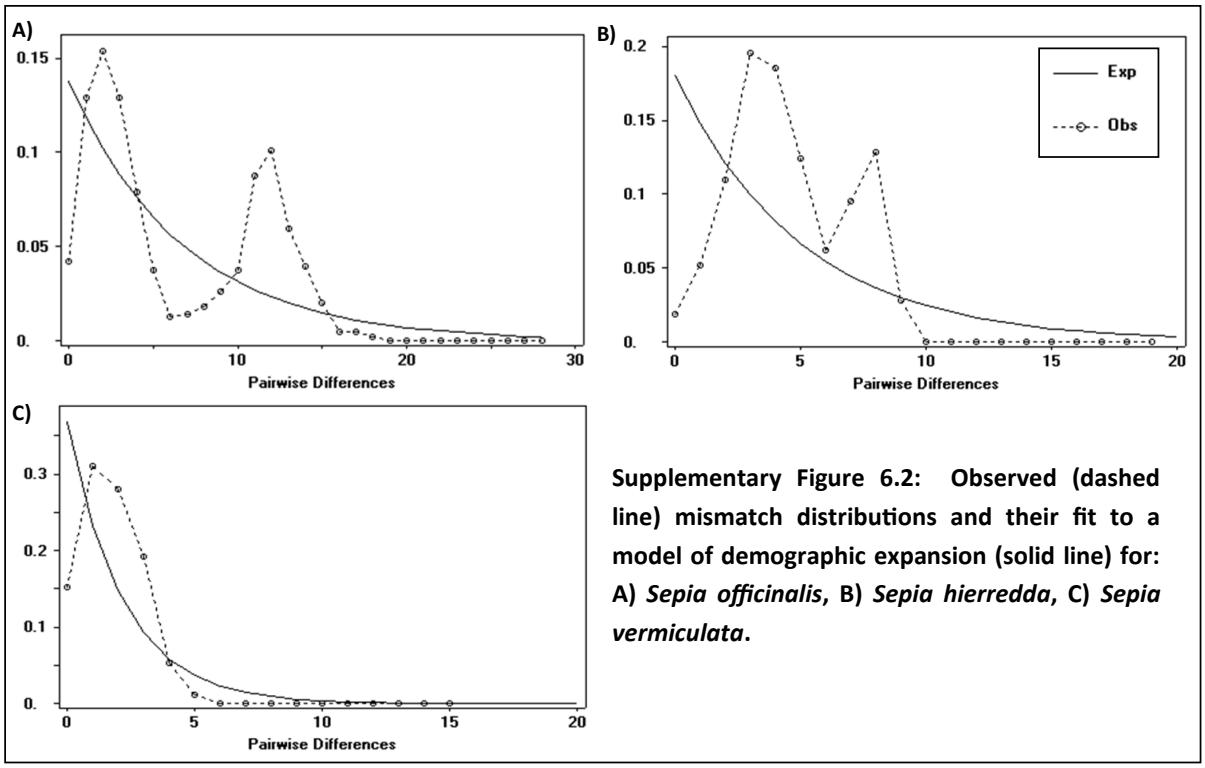
The demographic history of *S. officinalis*, *S. hierredda* and *S. vermiculata* was explored by conducting mismatch distributions in Arlequin v3.5.2.2. The sum of squared deviation (SSD; Schneider & Excoffier 1999) between observed and expected mismatch distributions, as well as Harpendings' (1994) Raggedness index were employed as test statistics for expansion models. For those samples that did not deviate significantly ($p>0.05$) from expectations under a model of sudden expansion, time of expansion was calculated using the formula $T = \tau/(2\ \mu)$, were μ= the mutation rate per site per year. A mutation rate of 2% per million years was employed as this is the most frequently used value for calibration of mtDNA phylogenies in cephalopods, including *Sepia officinalis* (Perez-Losada *et al.* 2007; Strugnall & Lindgren 2007; Ibáñez *et al.* 2011; 2012; Sales *et al.* 2017).

Results

Supplementary Table 6.1: Mismatch distribution values for *S. officinalis, S. hierredda* & *S. vermiculata* based on 839bp of concatenated COI & cytb. SSD=sum of squared distribution, τ= time since expansion expressed in mutational units, T_{EXP}= time of expansion, θ_0= population size before expansion, θ_1= population size after expansion. Significance ($p<0.05$) indicated in bold

	S. officinalis	*S. hierredda*	*S. vermiculata*
Mismatch mean	3.882	4.543	1.933
SSD	**0.009**	0.011	0.016
τ (95%CI)	2.844 (1.145-3.953)	3 (2.008-9.383)	2.305 (0.826-4.709)
Theta 0	0.012 (0.000-1.614)	1.900 (0.000-5.556)	0.011 (0.000-0.939)
Theta 1	195.625 (6.564-294.698)	3414.978 (11.833-147.642)	14.726 (2.216-945.980)
T_{EXP} (95% CI)	N/A	89,392 (59833-279588)	68,682 (24612-140315)

Supplementary Figure 6.1: Chronogram showing estimated divergence times between *Sepia officinalis*, *Sepia hierredda* and *Sepia vermiculata*, calculated using the RelTime method applied to the ML topology generated using 839bp of concatenated COI and cytb. Estimated node ages (MY) are given on the node, with bars representing the 95% confidence interval of age estimates. Timescale, in million years (MY) is provided on the x axis. Red circles correspond to fossil calibrations described in methods.



Supplementary Figure 6.2: Observed (dashed line) mismatch distributions and their fit to a model of demographic expansion (solid line) for: A) *Sepia officinalis*, B) *Sepia hierredda*, C) *Sepia vermiculata*.

Timing of divergence between and diversification within the *Sepia officinalis* species complex

Timetree estimates (Supplementary Figure 6.1) placed the divergence of *S. officinalis* from *S. hierredda* and *S. vermiculata* 7.27MYA (95%CI=12.38-2.15), with the split between west African *S. hierredda* and southern African *S. vermiculata* placed at 2.71MYA (95%CI=6.04-0.00). Within *S. officinalis*, intra-specific diversification was estimated as occurring 0.55MYA (95% CI=1.08-0.03) when the Mediterranean clade split from all other *S. officinalis*, followed by the divergence of the Portuguese clade 0.33MYA (95%CI=0.70-0). Within *S. hierredda*, intra-specific diversification was estimated as occurring 0.47MYA (95%CI=1.17-0) when Angolan populations diverged from the rest of tropical west Africa.

Demographic history of *S. officinalis*, *S. hierredda* and S*. vermiculata*

Mismatch distributions (Supplementary Figure 6.2) were unimodal and consistent with a model of demographic expansion for *S. vermiculata*, supported by non-significant values of SSD. Mismatch distributions for *S. officinalis* and *S. hierredda* were both bimodal, and inconsistent with a model of demographic expansion, supported by significant values of SSD for *S. officinalis*, however SSD for *S. hierredda* was non-significant, supporting a demographic expansion in this species (Supplementary Table 6.1). Using a 2% per million year mutation rate the demographic expansion of *S. vermiculata* was placed ~68,682YA (95%CI=140,315-24,612), and the demographic expansion of *S. hierredda* (although this not entirely supported), was placed ~89,392YA (95%CI= 279,588-59,833).

SUPPLEMENTARY MATERIAL FOR CHAPTER 7

7.1 Description of method applied to construct species distribution models

In order to assess the vulnerability of a number of commercially important species (for which genetic analysis reported a prolonged isolation of Angolan and/or South African populations/species) under predicted environmental changes, genetic data was used to inform Species Distribution Models (SDMs), forecast under various climate change scenarios. Post 1950s occurrence data was compiled from records obtained from the Global Biodiversity Information System (GBIF) as well as the Ocean Biogeographic Information System (OBIS), and occurrence data was partitioned so that each isolated genetic unit (e.g. Angola and South Africa) was modelled independently. ENMs were constructed in MAXENT v3.3.3 (Phillips & Dudik 2008) using the parameters outlined in chapter 5.1, however models were projected onto bioclimatic variables forecast to 2050 and 2100 under various climate change scenarios (e.g. RCP45, RCP60, RCP85) obtained from Bio-ORACLE (Tyberghein *et al.* 2012; Assis *et al.* 2018).